Department of Virology
Faculty of Medicine
University of Helsinki
Finland

NOVEL HUMAN PROTOPARVOVIRUSES: EPIDEMIOLOGY AND CLINICAL IMPACT

Elina Väisänen

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine, University of Helsinki, for public examination in Haartman Institute Lecture Room 2, Haartmaninkatu 3, Helsinki, on September 11th, 2020, at 12 noon.

Helsinki 2020

Supervisor

Maria Söderlund-Venermo
PhD, Docent
Department of Virology
Medicum
University of Helsinki
Helsinki, Finland

Reviewers

Carita Savolainen-Kopra
PhD, Docent
Expert Microbiology unit
National Institute for Health and Welfare
Helsinki, Finland

Sisko Tauriainen
PhD, Docent
Institute of Biomedicine
University of Turku
Turku, Finland

Opponent

Kristina Broliden,
Professor, Senior physician
Bioclinicum
Department of Medicine Solna
Karolinska Institutet
Stockholm, Sweden

ISBN 978-951-51-6407-0 (paperback)
ISBN 978-951-51-6408-7 (PDF, available at http://ethesis.helsinki.fi)

Helsinki University Printing House
Helsinki 2020

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

Once upon a time there was a tale of a ketchup bottle.
You open it and first nothing comes out.
But if you are patient and try long enough, the ketchup starts running.
And then there will be plenty.

But what the tale did not tell us
is that sometimes it is best to open another bottle.

CONTENTS

ABSTRACT

In the beginning of the 21st century, the development of next generation sequencing (NGS) methods revolutionized the discovery of novel viruses and laid the foundations of a new era in virology. With NGS, the amount and quality of sequence data increased tremendously, and the decreasing price made large-scale screening possible. As a result, hundreds of novel viruses have been identified, and especially the number of novel small DNA viruses, including parvoviruses, has grown substantially.

This thesis consists of studies of three novel parvoviruses, bufavirus (BuV), tusavirus (TuV), and cutavirus (CuV). All three viruses were originally identified by NGS from the diarrheal feces of children between 2012 and 2016, and they were the first viruses in the *Protoparvovirus* genus putatively infecting humans. Many of the animal viruses in this genus are known pathogens: for example, canine parvovirus (CPV) can cause severe gastroenteritis with high mortality among puppies. Currently, there are three genotypes of BuV and one genotype each of CuV and TuV. BuVs have been detected mainly in fecal samples, whereas CuV DNA has been detected in skin biopsies of cutaneous T-cell lymphoma (CTCL) and melanoma patients as well. Any findings of TuV have been rare. Here, we developed diagnostic methods for detecting viral DNA and IgG antibodies against BuV1-3, TuV and CuV, and analyzed 4300 samples from 11 cohorts on four continents to determine the epidemiology and potential clinical impact of these viruses.

The BuV DNA was detected in diarrheal fecal samples from both adults and children in Finland, although with low prevalence. In addition, one nasal swab harbored BuV DNA. These studies were among the first to show BuV circulation in Europe and among adults, and further, the nasal swab is still the only non-fecal human sample containing BuV DNA. Although we detected BuV exclusively in diarrheic feces, our results did not strongly support a major role of BuV in gastroenteritis.

One of the main findings in this thesis was the remarkable difference observed in the BuV seroprevalence between different populations: in the Middle East and Africa, the BuV IgG antibodies were detected in 56-85% of the adult population whereas in Finland and in the USA the seroprevalence was very low, <4%. This indicated that Iraq, Iran and Kenya are endemic areas for BuV. In addition to the seroprevalence difference, the predominant BuV type varied: it was BuV1 in the Middle East and BuV3 in Kenya. In Kenya, an age-dependent increase in BuV seroprevalence in children was apparent, and it seems that BuVs are not infecting young children in particular. The next important step in BuV studies will be identifying patients with acute primary infection to elucidate the symptoms and clinical picture of the infection. Our data on the geographical distribution of BuVs will help defining suitable locations for such studies.

In contrast to that of BuV, the CuV seroprevalence was low in all populations analyzed. Therefore, it is interesting that CuV DNA was detected significantly more often in the lesional skin biopsies from CTCL patients in Finland than in those of transplant patients

or healthy adults. This indicates that there is an association between the CuV DNA presence and CTCL, however, it is not known whether this association is causal, accidental or something else. The analysis of additional skin biopsies from the CuV DNA-positive patients revealed CuV DNA in every available skin sample, including both healthy and malignant tissue. Furthermore, lymph nodes harbored CuV DNA, while the prostate samples were CuV negative. Serological analysis of archived serum samples showed that the patients with CuV DNA in the skin had CuV IgG-antibodies already 5-21 years before the skin biopsies were taken. This suggests that CuV can persist for decades after primary infection similarly to human parvovirus B19. However, even if some preliminary disease associations exist, the role of CuV in CTCL or other skin cancers needs further investigations.

TuV findings, both DNA and antibodies, were absent or scarce. TuV DNA was not detected in any of the skin tissues, and TuV IgG was detected only in one child and one transplant patient among the serum samples of 1500 individuals ranging four continents. Overall, more studies are needed to confirm if TuV truly is a human virus or just accidentally occurring in human samples.

TIIVISTELMÄ

Vuosituhannen vaihteessa menetelmät syväsekvensointiin ja virusten rikastamiseen kehittyivät suurin harppauksin ja uusien virusten löytäminen erilaisista näytemateriaaleista helpottui merkittävästi. Tämä johti uuden aikakauden alkamiseen virologiassa, kun lukuisia uusia viruslajeja ja jopa aiemmin tuntemattomia virusheimoja tunnistettiin. Varsinkin pienten DNA-virusten, kuten parvovirusten, määrä on kasvanut merkittävästi.

Tässä väitöskirjassa tutkittiin kolmea uutta parvovirusta, bufavirusta (BuV), tusavirusta (TuV) ja cutavirusta (CuV), jotka kaikki löydettiin syväsekvensointi-menetelmillä lasten ripuliulosteesta vuosina 2012-2016. Ne olivat ensimmäisiä ihmisiä infektoivia viruksia *Protoparvovirus*-suvussa. Suvun muiden, eri eläinlajeja infektoivien virusten tiedetään aiheuttavan vaarallisiakin tauteja. Näihin kuuluu esimerkiksi koiran parvovirus, joka aiheuttaa vakavia ripulitauteja ja johon liittyy korkea kuolleisuus pennuilla. Tällä hetkellä BuV:sta tunnetaan kolme genotyyppiä ja CuV:sta ja TuV:sta molemmista yksi. BuV:ta on havaittu lähes yksinomaan ulostenäytteissä eri puolilla maailmaa, kun CuV DNA:ta on löydetty myös syöpäpotilaiden ihonäytteistä mm. kutaanista T-solulymfoomaa (cutaneous T-cell lymphoma, CTCL) ja melanoomaa sairastavilta potilailta. TuV löydökset ovat olleet harvinaisia. Tässä väitöskirjassa pystytimme diagnostisia menetelmiä BuV1-3-, TuV- ja CuV-DNA:n sekä vasta-aineiden havaitsemiseen ja analysoimme 4300 näytettä yhdestätoista eri kohortista neljältä eri mantereelta selvittääksemme näin uusien virusten epidemiologiaa ja kliinistä merkitystä.

Löysimme BuV DNA:ta suomalaisten ripulipotilaiden ulosteesta sekä aikuisilta että lapsilta, joskin esiintyvyys oli hyvin matala. Virusta löytyi myös yhdestä nenästä otetusta näytteestä. Vaikka tulokset eivät viitanneet siihen, että BuV aiheuttaisi gastroenteriittejä ainakaan laajassa mittakaavassa Suomessa, tutkimuksemme olivat ensimmäisiä, jotka osoittivat BuV:n kiertävän Euroopassa ja että BuV-DNA:ta löytyy myös aikuisten näytteistä. Lisäksi nenän pyyhkäisynäyte on edelleen ainoa näyte, jossa BuV DNA:ta on tavattu ihmisissä muualla kuin ulostenäytteessä.

Eräs väitöskirjan päätuloksista oli havainto BuV IgG-vasta-aineiden yleisyyden suuresta vaihtelusta eri väestöissä: Suomessa ja USA:ssa vain muutamalla prosentilla aikuisista oli vasta-aineita, kun Irakissa, Iranissa ja Keniassa jopa 85% aikuisväestöstä oli BuV-IgG-positiivisia. Lähi-Itä ja Afrikka näyttävätkin olevan BuV:n kotoperäistä aluetta, BuV1:n ollen yleisin virustyyppi Lähi-Idässä ja BuV3:n Keniassa. Lisäksi havaitsimme, että BuV-IgG-seroprevalenssi kasvoi lapsilla iän myötä ja tulos viittaa siihen, etteivät bufavirukset infektoi erityisesti pieniä lapsia, vaan kaiken ikäisiä. Seuraava tärkeä askel BuV-infektion tutkimuksissa on löytää potilaita, joilla on akuutti infektio, ja näin taudin oireita ja kliinistä kuvaa voidaan selvittää. Tuloksemme viruksen maantieteellisestä jakautumisesta helpottavat jatkotutkimusten suuntaamista alueille, joissa akuutteja infektioita todennäköisesti esiintyy.

Päinvastoin kuin BuV:n kohdalla, CuV:n IgG-seroprevalenssi oli matala kaikissa tutkituissa maissa ja eri potilasryhmissä. Siksi olikin mielenkiintoista, että löysimme CuV DNA:ta merkittävästi useammin CTCL-potilaiden syöpänäytteistä kuin elinsiirtopotilaiden tai terveiden aikuisten ihonäytteistä. Muita kudosnäytteitä analysoitiin kolmelta CuV DNA-positiiviselta potilaalta, ja CuV DNA:ta löytyi kaikista ihonäytteistä riippumatta siitä, oliko näyte terveestä vai sairaasta ihosta. Lisäksi CuV DNA:ta havaittiin imusolmukkeissa, mutta eturauhasnäytteet olivat negatiivisia. Arkistoituja seeruminäytteitä tutkimalla pystyimme osoittamaan, että CuV DNA-positiivisilta henkilöltä löytyi CuV IgG-vasta-aineita jopa 21 vuotta ennen ihonäytteen ottamista. Näin ollen on hyvin todennäköistä, että CuV on säilynyt elimistössä vuosikausia akuutin infektion jälkeen samankaltaisesti kuin ihmisen parvorokkovirus (B19V).

Erot CuV:n esiintyvyydessä eri potilasryhmien välillä olivat tilastollisesti merkittäviä ja tulos viittasi siihen, että CuV:lla voisi olla jokin rooli CTCL:ssä. Tulokset eivät kuitenkaan kerro, oliko CuV taudin syy, seuraus vai oliko virus sattumalta näissä kudoksissa. Tarvitaankin lisää tutkimuksia, jotta CuV:n merkitys CTCL:ssä ja muissa ihon syövissä selviää.

TuV löydökset olivat harvinaisia. Emme löytäneet TuV DNA:ta yhdestäkään ihokudoksesta ja vasta-aineita havaitsimme vain yhdellä lapsella ja yhdellä aikuisella elinsiirtopotilaalla. Näin ollen jää edelleen vahvistamatta, onko TuV oikeasti ihmisen virus vai löytyykö virusta vain satunnaisesti ihmisistä, joihin virus on joutunut vahingossa.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I **Väisänen E**, Kuisma I, Phan T, Delwart E, Lappalainen M, Tarkka E, Hedman K, Söderlund-Venermo M. Bufavirus in feces of patients with gastroenteritis, Finland. Emerg Infect Dis. 2014 Jun;20(6):1078-80. doi: 10.3201/eid2006.131674.

II **Väisänen E**, Paloniemi M, Kuisma I, Lithovius V, Kumar A, Franssila R, Ahmed K, Delwart E, Vesikari T, Hedman K, Söderlund-Venermo M. Epidemiology of two human protoparvoviruses, bufavirus and tusavirus. Sci Rep. 2016 Dec 14;6:39267. doi: 10.1038/srep39267.

III **Väisänen E**, Mohanraj U, Kinnunen PM, Jokelainen P, Al-Hello H, Barakat AM, Sadeghi M, Jalilian FA, Majlesi A, Masika M, Mwaengo D, Anzala O, Delwart E, Vapalahti O, Hedman K, Söderlund-Venermo M. Global distribution of human protoparvoviruses. Emerg Infect Dis. 2018 Jul;24(7):1292-1299. doi: 10.3201/eid2407.172128.

IV **Väisänen E**[a], Fu Y[a], Koskenmies S, Fyhrquist N, Wang Y, Keinonen A, Mäkisalo H, Väkevä L, Pitkänen S, Ranki A, Hedman K, Söderlund-Venermo M. Cutavirus DNA in malignant and nonmalignant skin of cutaneous T-cell lymphoma and organ transplant patients but not of healthy adults. Clin Infect Dis. 2019 May 17;68(11):1904-1910. doi: 10.1093/cid/ciy806.
[a], shared first authorship

The publications are referred to in the text by their roman numerals.

The publications I-III have been published in open access journals and the original electronic versions are freely available. The publication IV has been reproduced with permission from the respective copyright holders.

ABBREVIATIONS

aa	amino acid
AAP	assembly activating protein
AAV	adeno-associated virus
AGE	acute gastroenteritis
AIDS	acquired immune deficiency syndrome
ARTI	acute respiratory tract infection
B19V	human parvovirus B19
BHQ	Black Hole Quencher
BKV	human polyomavirus BK
bp	base pair
BPV	bovine parvovirus
BuV	human bufavirus
cap	capsid protein, used with AAVs instead of VP
CPV	canine parvovirus
CTCL	cutaneous T-cell lymphoma
CSN	central nervous system
CSF	cerebrospinal fluid
CuV	human cutavirus
DNA	deoxyribonucleic acid
EIA	enzyme immunoassay
EM	electron microscopy
EPC	erythroid progenitor cell
ETS	epitope-type specificity
FFPE	formalin-fixed, paraffin-embedded
GE	gastroenteritis
GT	genotype
H-1PV	rat protoparvovirus H-1
H5	HighFive insect cells
HAART	highly active antiretroviral therapy
HBoV	human bocavirus
HBoV1	human bocavirus 1
HBoV2	human bocavirus 2
HBoV3	human bocavirus 3
HBoV4	human bocavirus 4
HCV	hepatitis C virus
HIV	human immunodeficiency virus
IBD	inflammatory bowel disease
ICTV	International Committee on Taxonomy of Viruses
IgG	immunoglobulin G
IgM	immunoglobulin M
ISH	in situ hybridization

IU	international unit
JCV	human polyomavirus JC
kb	kilobase
kDa	kilodalton
MF	mycosis fungoides
mORF	putative short middle protein
MVC	minute virus of canines
MVM	minute virus of mice
mRNA	messenger ribonucleic acid
neg	negative
NGS	next generation sequencing
nm	nanometer
nPCR	nested PCR
NP1	nuclear phosphoprotein 1
NPS	nasopharyngeal sample
NS	non-structural gene or protein
NS1	non-structural protein 1
NS2	non-structural protein 2
nt	nucleotide
OD	optical density
ORF	open reading frame
p5	HBoV promoter
p6	B19V promoter
PARV4	human parvovirus 4
PCR	polymerase chain reaction
p.i.	post-infection
PLA2	phospholipase A2
pos	positive
PPV	porcine parvovirus
qPCR	quantitative polymerase chain reaction
rep	replication initiator protein, used with AAVs instead of NS
RNA	ribonucleic acid
SAT	small alternatively translated protein
SIV	simian immunodeficiency virus
ssDNA	single-stranded DNA
TuV	human tusavirus
tx	transplant
VLP	virus like particle
VP	virus (capsid) gene or protein
VP1	virus (capsid) protein 1
VP1u	VP1 unique N-terminal end
VP2	virus (capsid) protein 2
VP3	virus (capsid) protein 3
VR	variable region (in the capsid)

PREFACE

The development of next-generation sequencing (NGS) and metagenomics revolutionized the discovery of novel viruses and laid the foundations of a new era in virology. Whereas in the 20th century novel human viruses were mainly discovered by tedious, demanding and often biased methods such as cell culture or electron microscopy (EM), in the 21st century the sequence-independent amplification and NGS enabled detecting unknown viruses in a fast and straightforward manner. As a result, hundreds of novel viruses have been identified, and especially the number of novel small DNA viruses, including parvoviruses, has grown substantially. One common feature these small viruses share is that they grow poorly if at all in commonly used cell cultures. In addition, the capsid structure is often very stable and preserves the viral DNA both in the samples as well as during the virus enrichment steps often used prior to NGS.

The one big advantage cell culture has over NGS is the isolation of the novel virus itself. In comparison, when the virus is identified with NGS, it is only a sequence on a computer, and no knowledge of the infectivity, pathogenicity or structure of the virus is obtained. Neither is it known, whether the detected viral sequences are derived from an infection of the individual or if the sequence even is that of a true virus. Thus, NGS is only the first step in the process and further studies are needed to describe the virus, its biology, epidemiology as well as the clinical picture of the infection.

Regarding this thesis, the previous sentence sums up the purpose and aim of the project. The starting point was the novel putative human protoparvoviruses detected by NGS, and the thesis describes the diagnostic methods and the sample analysis that were carried out in order to elucidate the properties of these novel parvoviruses.

The literature review focuses on the overall properties of parvoviruses as well as on the specifics of the pathogenic human parvoviruses, human parvovirus B19 (B19V) and human bocaviruses (mainly HBoV1). The currently used parvovirus diagnostics is also described. An introduction to the novel human protoparvoviruses is given although more detailed information is available in the results and discussion part of this thesis.

1 REVIEW OF THE LITERATURE

1.1 OVERVIEW OF PARVOVIRUSES

Parvoviruses are small (21-28 nm), non-enveloped, single-stranded DNA (ssDNA) viruses that infect various animal species ranging from insects, crustaceans and starfishes to humans and numerous other mammals. The name "parvo" derives from the Latin word *parvus* meaning small, and it was introduced in 1966 to describe these small DNA viruses[1]. In general, parvoviruses are host specific and require rapidly dividing cells as the short 4-6 kb genome codes only for a few major proteins: non-structural proteins (NS or rep) on the left side and structural proteins (VP or cap) on the right side of the genome[2]. Although parvoviruses rely heavily on the host cell machinery and utilize for example the host DNA polymerase(s), these viruses replicate autonomously, except the members of the *Dependoparvovirus*, which need a helper virus.

Taxonomy of the parvoviruses. The parvovirus study group within the International Committee on Taxonomy of Viruses (ICTV) has set up demarcation criteria for distinguishing the viruses in the *Parvoviridae* family further into subfamilies, genera and species, and currently, in January 2020, the classification is made based on the NS1 sequence. The two sub-families, *Parvovirinae* and *Densovirinae*[2], are divided primarily based on the host that the viruses infect (vertebrates vs. invertebrates, respectively), although this distinction is largely, but not perfectly, supported by the phylogeny of the NS1 protein sequence. A parvovirus genus is formed when the viruses are monophyletic and share ≥30% of the NS1 amino acid (aa) sequence, while viruses that share >85% of the NS1 aa sequence are considered to be one species[2]. Based on these criteria, the two parvoviral subfamilies are divided into 13 genera with over 75 species (in Jan 2020).

The *Parvovirinae* subfamily consists of eight genera (Fig. 1), and human infecting parvoviruses have been detected in five: *Erythroparvovirus* (human parvovirus B19 [B19V]), *Dependoparvovirus* (adeno-associated viruses [AAVs]), *Bocaparvovirus* (human bocaviruses [HBoV1-4]), *Tetraparvovirus* (human parvovirus 4 [PARV4]), and *Protoparvovirus* (bufavirus [BuV], tusavirus [TuV] and cutavirus [CuV])[2]. The subfamily *Densovirinae* is divided into five genera: *Ambidensovirus, Iteradensovirus, Brevidensovirus, Hepandensovirus,* and *Penstyldensovirus*[2].

Until recently, all identified parvoviruses have followed the "rule" that viruses within *Parvovirinae* infect only vertebrates and viruses in *Densovirinae* infect invertebrates. However, a rapidly growing number of distinct parvoviruses have been identified that infect vertebrates, e.g. reptiles, mammals, and birds, but cluster phylogenetically with densoviruses, and therefore, the formation of a third subfamily* has been suggested[2,3].

*In March 2020 the ICTV accepted the revised taxonomy for the family *Parvoviridae* including a third subfamily, named *Hamaparvovirinae*. Pénzes et al. Arch Virol. 2020 Jun 12. doi: 10.1007/s00705-020-04632-4

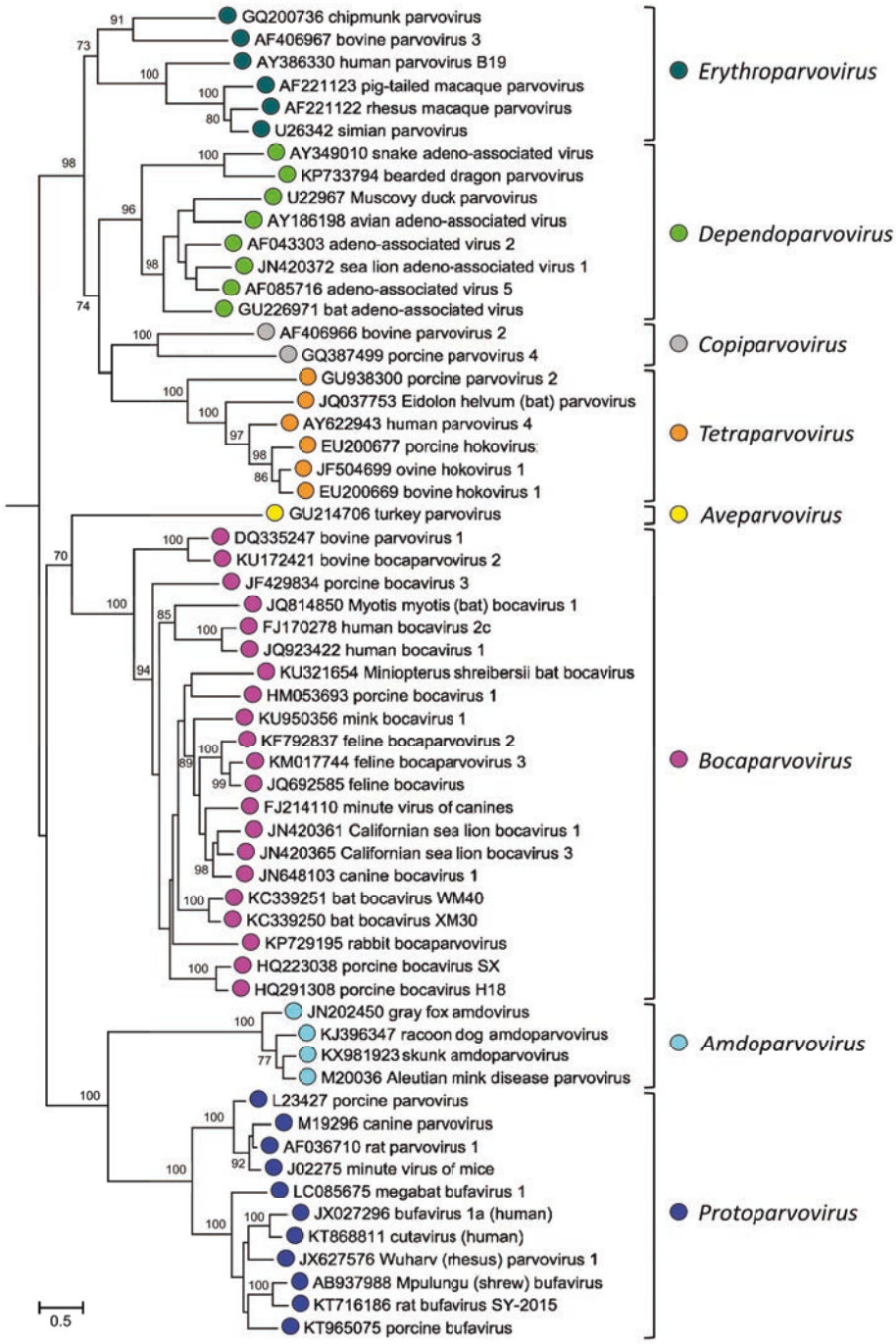


Figure 1. **Phylogenetic relationships of the different genera within the *Parvoviridae* subfamily.** The phylogeny is based on NS1 amino acid sequence and follows the ICTV demarcation criteria valid on January 2020. Reproduced and modified from Cotmore et al.[2] with permission.

Genome structure. The short ssDNA genome of parvoviruses consists of two major open reading frames (ORFs), NS and VP, and at both ends of the genome there are terminal hairpin structures, which are either homo- or heterotelomeric, i.e. identical or different, depending on the species (Fig. 2). Most parvoviruses code also for smaller ancillary proteins that utilize usually alternate or overlapping ORFs for transcription. Such include nucleoprotein 1 (NP1) of the ave- and bocaparvoviruses[4], assembly activating protein (AAP) of the dependoparvoviruses[5] and the small alternatively translated protein (SAT)[6] and NS2 of the amdo- and protoparvoviruses[7–9] (Fig. 2).

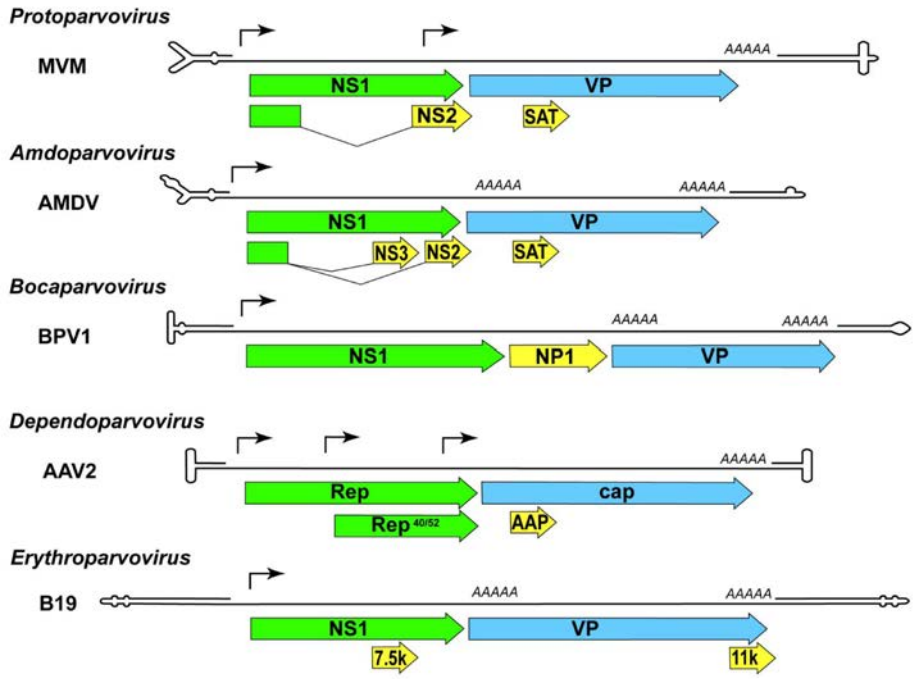


Figure 2. Genome structure of parvoviruses in different genera within the *Parvovirinae* subfamily. Genomes of the type species of each genus are shown as a single line with terminal hairpin structures. The hairpins are scaled approximately 20× relative to the rest of the genome. The ORFs encoding proteins are illustrated with boxed arrows: the green arrows represent nonstructural NS proteins, the blue arrows the capsid forming VP proteins and the yellow arrows the smaller ancillary proteins. Transcriptional promoters are indicated by solid arrows and polyadenylation sites by the AAAAA. AAV, adeno-associated virus; B19, human parvovirus B19; BPV, bovine bocavirus; MVM, minute virus of mice. Reproduced from Cotmore and Tattersall 2014[10] with permission.

The NS1 is a non-structural replication initiator protein that has both endonuclease and DNA helicase domains, whereas the roles of the often genus-specific ancillary proteins are versatile and not fully elucidated for all viruses. NS1 is considered most conserved as it is indispensable for the viral replication, and therefore the phylogenetic

classification of parvoviruses is made based on the NS1 sequences[2]. The parvovirus capsid protein is coded by a single gene in the right-hand side of the genome that includes the entire VP1 sequence. The same ORF encodes also shorter VP2 or VP3 proteins, which are truncated forms of VP1 with identical C-terminal sequence. With most parvoviruses the VP1 unique N-terminal end (VP1u) contains a phospholipase A2 (PLA2) domain, a calcium binding loop, and a nuclear localization signal.

Replication. In general, parvoviruses need rapidly dividing cells for replication as they cannot drive the cells into S phase. This also limits the cell types parvoviruses can replicate in, and many parvoviruses have very specific host cells in which they optimally grow in. The replication is carried out in the nucleus with the rolling hairpin mechanism utilizing the terminal hairpins as primers[10,11], and the single-stranded genomes are packaged into the ready-made capsid. Parvoviruses with identical terminal repeats, for example B19V (Fig. 2), encapsidate both minus and plus strands in equal moiety (in separate capsids), whereas the viruses with heterotelomeric hairpins, for example minute virus of mice (MVM, Fig. 2), encapsidate predominantly the negative sense strand[10]. All virions, regardless of the polarity of the packaged strand, are considered infectious, however, this not been studied.

Structure. The capsid consists of 60 copies of the VP proteins, which form a structure with T=1 icosahedral symmetry. The major capsid protein is the shortest, VP2 or VP3, and usually there are only 5-10 copies of the VP1 in the native viral capsid[10,12,13]. In vitro the shorter proteins can form the capsid alone, whereas VP1 cannot, and this shows that the C-terminus is responsible for the capsid assembly[14]. The mature viral particle does not contain any lipids or carbohydrates, and therefore the structure is very stable: parvoviruses tolerate elevated temperatures as well as a wide range of pH conditions, and they are resistant to many disinfectants[2,15].

The capsid structure of approximately 100 parvoviral particles have been published, including the X-ray crystallographic and cryo-EM studies with both capsid structures alone and together with receptors or antibodies[13]. In all defined structures, the VPs form a highly conserved core of an eight-stranded anti-parallel β-barrel motif (Fig. 3). Another conserved structure is an α-helix located between βC and βD strands (Fig. 3). The surface loops are located between the β-strands, and these loops contain the variable regions (VRs), which have the highest sequence and structural diversity between parvoviruses.

The 60-mer capsid is assembled with 2-, 3-, and 5-fold symmetry-related interactions (Fig. 4). The icosahedral 2-fold axis has a preserved depression among parvoviruses, while the 3-fold protrusions vary. At the 5-fold axis, five β-barrels form a cylindrical channel that has been proposed to be utilized for VP1 externalization during infection as well as for genome packaging.

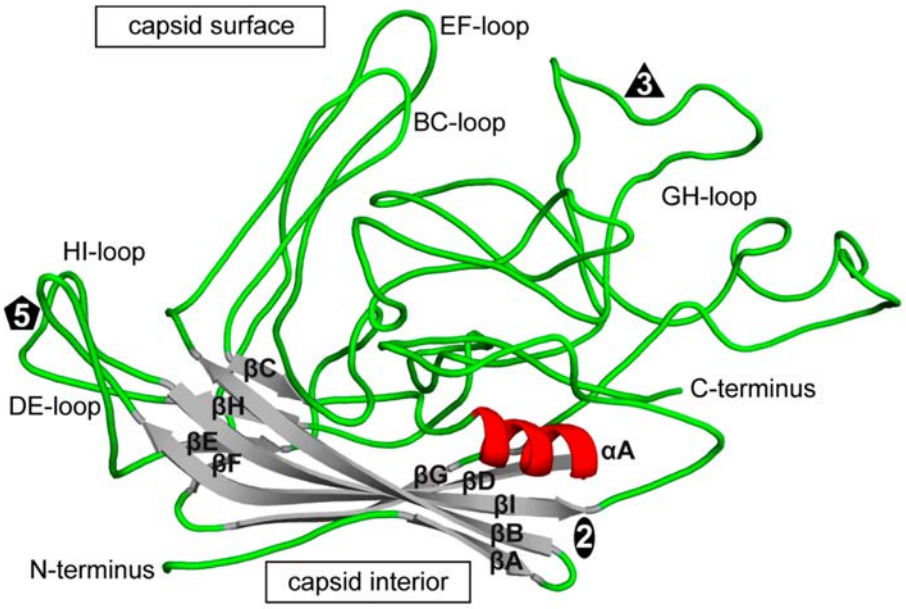


Figure 3. **Common features of parvovirus VP monomer structure exemplified by canine parvovirus (CPV) VP monomer.** The grey box arrows depict the eight β-strands in the conserved β-barrel motif (labeled βA to βI), and the red serpentine depicts the conserved α-helix. Surface loops contain the highest variability between different parvoviruses (see Fig. 8 for some examples of the loop differences between selected parvoviruses). The icosahedral axis locations for 2-fold (oval), 3-fold (triangle), and 5-fold (pentagon) as well as N- and C-termini are indicated (see Fig. 4 for a symmetry diagram illustrating the positions of the icosahedral symmetry axes on the capsid surfaces). Reproduced from Mietzsch, Pénzes and Agbandje-McKenna 2019[13] with permission.

The structural studies have demonstrated that the capsid structures are more conserved than the protein sequences would suggest, and that even the variable regions have similar topologies among parvoviruses (Fig. 4)[13,16,17]. For example, B19V and minute virus of mice (MVMp) share only 14.5% VP2 sequence identity, whereas the structural similarity is 41.5%[13]. In the defined structures, only capsid forming part of the VP (VP2 or VP3) has been resolved, whereas the structure of the flexible VP1 N-termini (VP1u) is unknown. With most parvoviruses the VP1u is believed to be positioned inside the capsid and not affect the capsid structure. The flexibility is associated to the cluster of glycines present close to the N-terminal end of the major VP, and this disordered glycine-rich region is present among most species in the *Parvovirinae* subfamily.

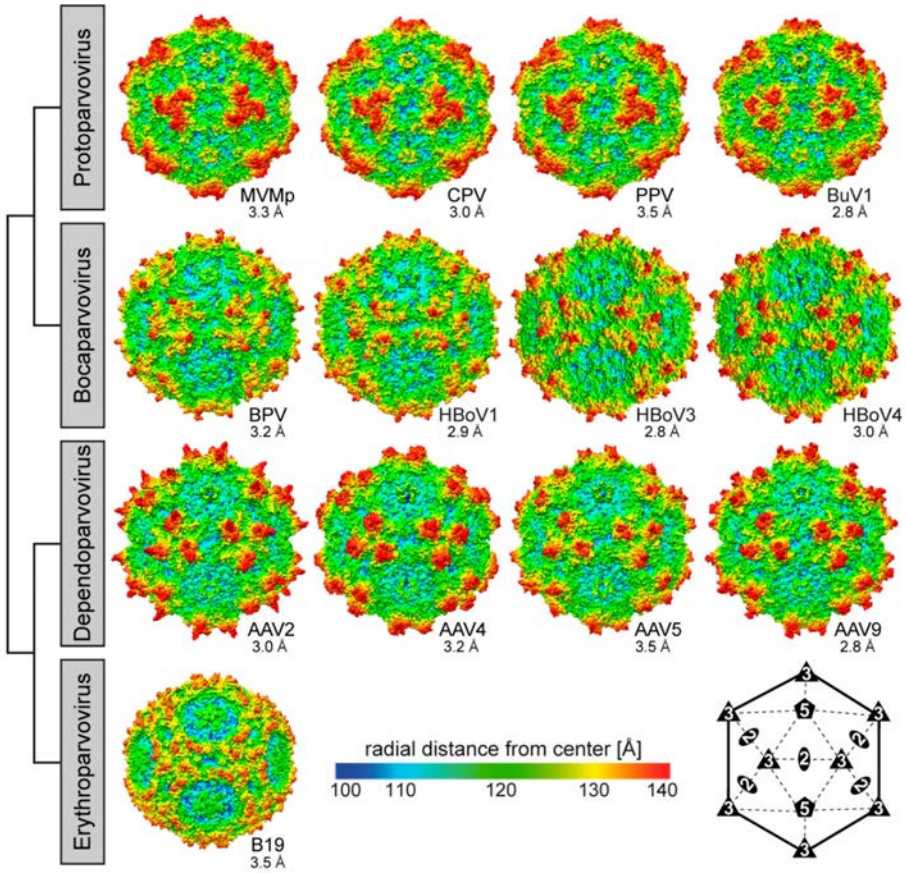


Figure 4. **Capsid structures of selected parvoviruses in the *Parvovirinae* subfamily.** The capsids are viewed from the 2-fold axes and colored according to radial distance as indicated by the scale bar (from blue [inside] to red [outmost parts]). AAV, adeno-associated virus; B19, human parvovirus B19; BPV, bovine bocavirus; BuV, human bufavirus; CPV, canine parvovirus; HBoV, human bocavirus; MVMp, minute virus of mice; PPV, porcine parvovirus. Reproduced from Mietzsch, Pénzes and Agbandje-McKenna 2019[13] with permission.

1.2 HUMAN PARVOVIRUSES

1.2.1 HISTORY OF HUMAN PARVOVIRUSES

The story of parvoviruses started in 1959, when Kilham and Olivier published their findings considering a small and stable virus isolated from experimental rat tumors[18]. The virus was later named Kilham rat virus, and during the next decade, several related viruses were isolated from different laboratory rodents. In 1965 similar small viruses were detected with electron microscopy as contaminants of a rhesus monkey kidney cell culture infected with the simian adenovirus SV15[19]. Although these particles were defective viruses and capable of replication only in the presence of simian or human adenoviruses, they were believed to be true viruses[19]. The virus was named adeno-associated virus (AAV), and it is considered the first identified human parvovirus.

Ten years later in 1975 human parvovirus B19 (B19V) was discovered[20]. As human parvoviruses do not grow easily or at all in the commonly used cell culture systems or in animal models, it is not surprising that the discovery of B19V was a coincidence like the discovery of AAV: B19V caused a false reaction in a hepatitis B virus antigen test in blood-donor screening[20]. Whereas AAV is in general regarded as non-pathogenic, B19V has been shown to cause several diseases and symptoms such as a childhood rash disease *erythema infectiosum* and acute joint pain similar to that of arthralgia and arthritis. More severe complications include chronic anemia in the immunocompromised, and miscarriage and fetal death during pregnancy. During the next three decades no other human-infecting parvoviruses were identified, and thus it was believed that B19V is the sole pathogenic human parvovirus.

However, that view changed in 2005. In the early 2000s, novel methods for large-scale screening of previously unidentified viruses were developed, which facilitated the start of active virus hunting[21]. These methods included virus enrichment and sequence-independent amplification followed by cloning and sequencing[22,23], although the last steps were fairly soon substituted by next-generation sequencing and metagenomics. Using these techniques, two research groups independently described two novel parvoviruses in human samples in 2005: i) human parvovirus 4 (PARV4) was identified in a serum sample of an intravenous drug user with acute infection symptoms, and ii) human bocavirus 1 (HBoV1) was identified in nasopharyngeal aspirates of children with respiratory symptoms[23,24].

When the first sequence of a novel virus is described, the closely related viruses are often detected soon after. This happened also with PARV4 and HBoV. Another PARV4 type, PARV4 genotype 2 (originally named PARV5), was detected in human plasma pools in 2006 and a third genotype in bone marrow and lymph nodes in 2008[25,26]. For HBoV, three additional HBoVs, HBoV2-4, were discovered from fecal samples of diarrheic children between 2009 and 2010[27–29]. However, whereas the PARV4 genotype sequences resemble each other closely (only ~8% divergence on the nt level and ~3% on aa[30–32]), and are thus genotypes, HBoV1-4 differ more and form phylogenetically two species:

HBoV1 and HBoV3 as well as gorilla bocavirus belong to one species and HBoV2 and 4 to the other species[2,33]. Currently HBoV1 is established as the causative agent for respiratory tract infections. PARV4, on the other hand, has been studied for 15 years, but the clinical relevance is still unknown (reviewed in [31]).

Between 2012 and 2016 the trend of finding novel human parvoviruses continued, and three more parvoviruses were identified by NGS: bufavirus (BuV), tusavirus (TuV), cutavirus (CuV)[34–36]. All of them were detected in diarrheal feces of children, and they were the first viruses putatively infecting humans in the *Protoparvovirus* genus. This thesis describes the studies of BuV, TuV and CuV, and these viruses will be discussed in detail in forthcoming chapters.

1.2.2 HUMAN PARVOVIRUS B19

B19V is the most widely known and studied human-pathogenic parvovirus, and when clinicians or laymen talk about parvoviral infections in humans, they usually mean B19V infections.

Genome structure and replication. The 5.6-kb ssDNA genome of B19V is very typical of parvoviruses and the long inverted terminal repeats are identical to each other (Fig. 2)[37,38]. Both minus and plus strands are encapsidated in equal numbers, i.e. half of the mature virions contain minus strand and the other half plus strand[38]. The B19V genome codes for non-structural NS1, capsid proteins VP1 and 2, as well as two short proteins, 7.5 kDa and 11 kDa (most recent transcription map is presented in [39]), of which the latter has been shown to have a regulatory role in VP expression as well as in apoptosis[40]. The protein expression is controlled by a single promoter, p6, which is located between the left terminal repeat and *NS1*. The virus capsid consists of >90% of VP2 and <10% of VP1[12], and the structure is similar to that of other parvoviruses (Figs. 4 and 8). Globoside (P-antigen) has been considered to be the primary receptor for B19V with the α5β1 integrin and Ku80 as possible co-receptors[41–43], however, a recent study showed that globoside is not needed during B19V cell entry[44]. Instead, globoside was crucial for the NS1 transcription, and consequently the lack of globoside blocked the B19V genome replication and capsid protein expression, which explains the resistance to B19V infection in the individuals lacking globoside[44,45].

Overall, B19V has a very narrow host cell tropism, it replicates only in the erythroid progenitor cells (EPCs) in the bone marrow[46,47]. Besides the EPCs, B19V can enter several other cell types, however, productive infection does not take place. This entry can be helped with antibodies, and the uptake has been shown occur through antibody-dependent enhancement (ADE) in monocytes, endothelial cells, and B cells[48–50].

Interestingly, after the primary infection, B19V DNA remains in tissues, for example in skin and in synovia[51–53], and can be detected with sensitive PCR methods for decades or even life long after the acute infection[53]. The mechanism for the persistence is not

known, however, long-lived B cells have been suggested be one cellular site for the tissue persistence[50].

Genotypes of B19V. The B19V strains can be divided into three genotypes (GTs), which differ from each other approximately 10% in the coding region (at the nt level)[54,55]. GT1 is the prototypic virus and is currently the predominant virus circulating worldwide[56–59]. The two other genotypes are rare: GT2 is mainly found in tissues and only sporadically in blood[53,55,60,61], and GT3 has been described to be endemic in Ghana and Brazil and occasionally found in India and Turkey[57–59,61]. Interestingly, based on the GTs detected in skin and other tissues of people from different age groups, the data suggest that GT1 replaced GT2 as the main circulating GT in the 1970s[50,53,62]: In the study by Norja *et al.*, GT1 DNA was detected in skin samples of all age groups, whereas GT2 DNA was present only in persons born 1973 or earlier[53]. The change in the GT was further shown in a study by Toppinen *et al.*, in which the remains of putative Finnish casualties from World War II soldiers harbored B19V genotypes 2 and 3, but not of genotype 1[62].

Clinically or serologically there is no difference between the three GTs. The symptoms described for GT2 and GT3 infections are identical to GT1, and in addition, cell culture experiments have shown similar results[63]. Serologically the three genotypes have been cross-reactive[63–65], although in one study a commercial IgG EIA assay utilizing GT1 antigen failed to detect a portion of GT3 IgG-positive samples[57].

Clinical aspects. B19V has been shown to be the etiological agent in several diseases and the causality has been established in various, well-defined studies, including experimental B19V infections of healthy adult volunteers[66].

B19V is transmitted mainly through the respiratory route, but infection can be acquired also through contaminated blood products[67,68] or transplacentally from mother to the fetus[69]. During acute infection, B19V DNA is first detected in the upper airways and then a high-titre viremia (>1E+10 viral copies per ml serum) follows[66]. The viral load then rapidly decreases; however, a low-level viremia can persist for months or years. B19V infection elicits a strong humoral response that starts with IgM antibody 1-2 weeks post-infections (p.i.) and is followed by IgG (Fig. 5)[66]. IgM is responsible for clearing up the viremia and they last up to six months. IgG, on the other hand, prevails and provides life-long immunity to B19V re-infection.

The most common disease caused by B19V is a childhood rash disease, called fifth disease or *erythema infectiosum*[70], although most infections are asymptomatic or mild. In adults, arthritis is a common symptom of acute B19V infection[71]. In the study of experimental B19V infection of healthy adults, two episodes with different symptoms were observed: during high viremia (1-2 weeks p.i) fever and other non-specific symptoms such as headache and itching occurred, and rash and arthritis were observed later (3 weeks p.i.), when the viremia had already been cleared and the antibody response developed (Fig. 5)[66].

More severe symptoms of B19V infection are seen in patients with immunosuppression or hematological diseases: persistent B19V infections may lead to chronic anemia in immunocompromised patients[72,73] and transient aplastic crisis may occur in patients with hematological disorders such as sickle cell disease[74,75].

The most devastating outcome of B19V infection can occur during pregnancy, when a primary infection of the mother may result in hydrops fetalis and miscarriage or fetal death, especially if the infection takes place during the first half of the pregnancy. The seroconversion rate among pregnant women has been estimated to vary from 1.5% up to 13% during epidemic year[76–79]. However, transmission of virus to the fetus has been estimated to occur in less than half of the cases[80,81], and fetal loss is even less likely (<10%)[81,82]. Moreover, if the baby is born alive, congenital abnormalities due to B19V are rare[83,84].

Occasionally B19V infection can lead to other severe diseases, such as pediatric myocarditis, although these cases are rare. In addition, many studies have linked B19V to various syndromes ranging from hepatitis and encephalitis to dermatological conditions and chronic fatigue, however, the results are often controversial, strong evidence is lacking, and in many cases the presence of B19V DNA due to the tissue persistence after primary infection is not accounted for. For the interested reader, B19V related diseases have been recently reviewed by Qiu et al.[39].

1.2.3 HUMAN BOCAVIRUS 1-4

Human bocaviruses 1-4 were discovered between 2005 and 2010, and the sequences showed similarity to the previously known bovine parvovirus (BPV) and minute virus of canines (MVC) in the *Bocaparvovirus* genus. Soon after the discovery, evidence of the role of HBoV1 in respiratory infections of children started to accumulate, while the role of HBoV2-4 in still unknown.

Genome structure and replication. One full-length genome of HBoV1 has been sequenced: it was 5543 nts long with non-identical hairpin structures in the ends of the genome. Whereas B19V encapsidates both strands in equal amounts, 95% of the bocavirus capsids contain the negative-sense genome[85]. HBoV utilizes one promoter, p5, and the genome codes for four NS proteins (NS1-4), a unique NP1 protein, and three VP proteins (VP1-3) of which VP3 is the major capsid protein[39,86–89].

HBoV1 infects airway epithelial cells, and cell culture experiments have demonstrated that HBoV1 can infect and replicate in the terminally differentiated cells[90]. Instead of using the DNA-polymerases of cell DNA-replication, HBoV1 may be replicated by the cellular DNA-repair DNA-polymerases, and thus the replication is cell cycle independent[90]. The permissive cells or replication strategies for HBoV2-4 are not known.

Besides the genomic ssDNA form, HBoV1-4 DNA has been detected in episomal forms[91–94]. Whether these circular dsDNA molecules are replication intermediates, persisting forms of HBoV DNA or normal forms of enteric HBoV2-4 strains, is not known.

Taxonomy of the four HBoVs. The NS1 and VP1/2 amino acid sequences of the four HBoVs differ from each other 10.8-26.8% and 8.5-19.9%, respectively[29], which is clearly more than what the B19V genotypes differ from each other. The divergence of NS1 has further consequences taxonomically: HBoV1 and HBoV3 as well as gorilla bocavirus belong to *Primate bocaparvovirus 1* and HBoV2 and 4 to *Primate bocaparvovirus 2*[2,33]. However, as the species classification does not follow the observed clinical picture of HBoV infections nor the serological cross-reactivity between the four HBoVs, the species naming has not been adopted in use[21]. In addition, HBoV3 clusters differently in the NS and NP1 regions compared to the capsid protein region: on NS and NP, HBoV1 is the closest relative of HBoV3, but in contrast, on VP, HBoV3 clusters with HBoV2 and HBoV4. This separation was further backed up with full genome analyses, and therefore it was suggested that HBoV3 could have originated from a recombination event between HBoV1 and HBoV2[29]. Similarly, HBoV4 was speculated to be a result from a HBoV2 and HBoV3 recombination event, however, the analysis was based on a single HBoV4 sequence[29].

Clinical aspects. Currently, HBoV1 is established as the causative agent of respiratory tract infections, whereas HBoV2-4 are considered to have a more enteric role. HBoV1-3 are commonly detected, whereas HBoV4 is very rare (recently reviewed in [95]).

The exact routes of infection of HBoV1-4 are not known, however, it has been hypothesized that HBoV1, a respiratory pathogen, would be transmitted through the respiratory route, whereas the possibly enteric HBoV2-4 would most likely utilize the fecal-oral route[95]. In comparison to B19V, HBoVs have not been considered to be threats during pregnancy[96,97], which is in line with the fact that nearly all adults are seropositive and thus primary infections among pregnant women are rare. However, in a recent study, HBoV DNA was detected in aborted fetal tissue and placenta, but whether the detected DNA derived from an infection during pregnancy, could not be determined[98]. In addition, discrepant results have been published of whether HBoVs present in blood products are a cause for concern[25,99,100].

HBoV1 infects mainly young children. During acute infection, the viral load in nasopharyngeal samples (NPS) is high and the patients are viremic. After the acute phase the viral DNA persists in the airways from weeks up to a year, and can thus be detected in healthy children as well[101]. As a consequence, it has been estimated that only a quarter of the children with HBoV1 DNA detected in the NPS actually have an acute HBoV1 infection[102,103]. HBoV2 and HBoV3 are detected mainly in fecal samples of children with or without gastroenteritis, but viremia can occur.

IgM and IgG antibodies are elicited during HBoV infection, and the IgG antibody response is long-lasting, but sometimes fluctuating[103,104]. HBoV1 seems to induce a stronger IgG response than HBoV2 and 3[103]. The HBoV strains are highly cross-reactive, which makes the analysis of virus-specific antibodies and serodiagnosis of an acute infection more difficult. The correct analysis therefore requires the use of a competition assay[105]. In addition, a phenomenon called “original antigenic sin” has been observed in

HBoV infection: when a child is first infected with one HBoV strain, the secondary infection with another HBoV strain does not induce specific antibodies towards the second strain, but instead, the IgG of the primary HBoV strain is boosted[106]. How this phenomenon affects the clinical picture of the secondary HBoV infections, is not known. Nevertheless, HBoV infections are very common, which is demonstrated by the prevalence of IgG antibodies in the population: over 90% of the less than 3 month-old babies have been shown to have maternal IgG antibodies[107], and by the age of six years 90-100% of the children had IgG of one of the HBoVs of which HBoV1 IgG is the most prevalent[103].

When using accurate diagnostic methods (see HBoV diagnostics), HBoV1 has been shown to be an important respiratory pathogen in young children. The symptoms are very similar to that of other respiratory viruses including common cold, cough, fever, rhinitis, bronchiolitis, diarrhea, and wheezing. Besides the mild infections, HBoV1 can cause severe lower respiratory tract infections and pneumonia requiring hospitalization, and life-threatening cases or even death have been reported with HBoV1 as the sole infectious agent[95,108–114].

1.2.4 DIAGNOSIS OF HUMAN PARVOVIRUS INFECTIONS

In virus diagnostics one of the main questions has been the rationale of the entire testing, if a potentially expensive test does not provide clinically useful information for the clinician. However, nowadays there is a growing number of antiviral drugs available, requiring the knowledge of the causative agent. In addition, the use of antibiotics can be reduced by knowing that the infection is caused by a virus instead of a bacterium, and accurate screening and diagnostics can also help controlling epidemics and directing vaccine campaigns when a vaccine-preventable virus is causing an outbreak[115]. One growing patient group, which benefit from the diagnosis of a wide range of virus infections, are the immunocompromised individuals. For these patients normally less or non-pathogenic viruses may become deadly, for example polyomaviruses BKV and JCV in the case of transplant patients[116,117].

In parvovirus diagnostics, serological testing of the patient's serum is one of the main approaches for achieving a useful diagnosis. Using IgM and IgG analyses the timing of the infection can be deduced, which is often critical, for example in pregnant women with B19V infection. Nucleic acid testing by quantitating the viral DNA or by measuring RNA can also be used to pinpoint an acute infection, although prolonged parvovirus shedding and viral DNA persistence complicate the interpretation of the PCR result.

To date, all human parvoviruses have been shown to induce antibody responses in humans, and therefore serological assays could be useful tools to study the infections and epidemiology of novel parvoviruses too. As the viremic phase, or the presence of viral DNA in specific samples, might not correlate with the symptoms of the patients or with the true acute infection, as seen with B19V and HBoV1, the results obtained from using

solely PCR might not be conclusive. The course of a typical parvovirus infection with viremia and antibody response is presented schematically in Figure 5.

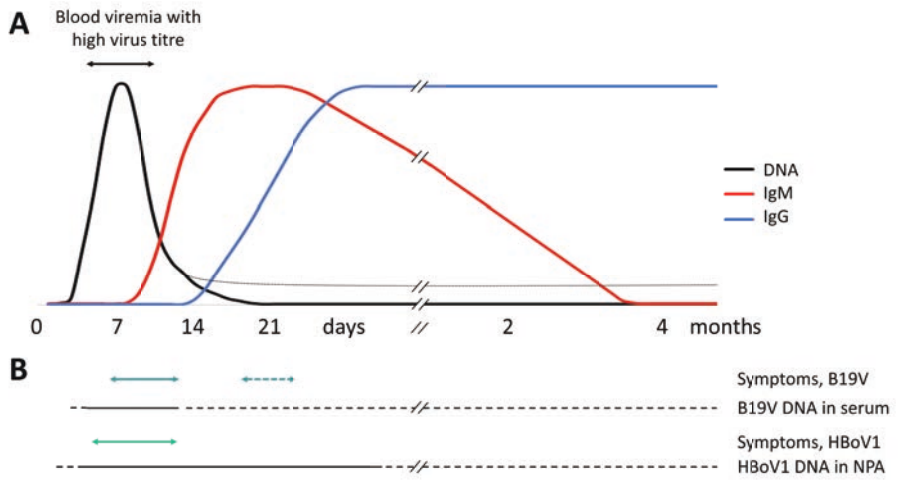


Figure 5. **Schematic view of typical human parvovirus infections.** (A) depicts viremia followed by antibody response, and (B) timing of symptoms and time frame of detecting DNA in B19V and HBoV1 infections. See text for further information. NPS, nasopharyngeal swab.

B19V diagnostics. Acute infections of B19V are diagnosed mainly by analyzing the patient's serum sample for IgM and IgG antibodies. A high B19V DNA quantity in the serum (>10E+06 copies per ml) also indicates acute infection, however, the viremic phase with high titres usually precedes the symptoms and often the sample is taken too late. In addition, after primary infection, B19V DNA persists in tissues and a low-level B19V viremia can be detected in immunocompetent individuals[118–120]. Thus, the DNA-positivity of a tissue or serum alone, especially when analyzed by qualitative PCR or when qPCR shows a low viral load, is not a good marker for an acute infection. With immunocompromised patients, the DNA testing might be the only option due to lack of antibody formation, and with pregnant women, PCR coupled with antibody testing can help in making the diagnosis.

The most reliable indicators for acute B19V infection are detection of IgM and seroconversion or a 4-fold increase of IgG in paired serum samples. When only one serum sample is available, the timing of the infection can be analyzed with IgG avidity, i.e. the functional affinity of the IgG, or with epitope-type specificity (ETS) assays[121–124]. Both of these assays utilize the different properties of early and matured IgG against B19V: in the acute infection IgG antibodies to linear VP2 epitopes (peptides or linearized VP2) can be detected and the IgG affinity to conformational VP2 (capsid) is weak, whereas later on the IgG to linear epitopes disappear and the affinity of the IgG to conformational epitopes increases.

For ensuring the safety of blood products, the authorities in the USA have implemented regulations of the B19V-DNA load in plasma pools not to exceed 1E+04 IU/ml (https://www.fda.gov/vaccines-blood-biologics/biologics-guidances/blood-guidances and [125,126]). In Europe similar recommendation apply. The cut-off was not set lower as the neutralizing B19V antibodies in the plasma pools are considered to protect the seronegative recipients from infection.

HBoV1-4 diagnostics. Diagnosing acute HBoV1 infections are more complicated than for B19V infections due to several reasons. Firstly, HBoV1 DNA can be detected in nasal secretions for months after the acute infection, and thus healthy individuals have HBoV1 DNA in their samples[101]. HBoV1 has also been detected in fecal samples, however, it is not known whether HBoV1 caused GE or if the detected virus originates from the nasopharynx[127,128]. Secondly, the four HBoVs have strong cross-reactivity in serological assays, and it has been shown that prior immunity to one HBoV may hamper or even prevent the antibody response towards another HBoV[103,106]. Therefore, it is critical to use proper methods for HBoV primary infection diagnosis, and it is recommended that HBoV1 acute primary infection diagnosis should include at least two of the following: by qPCR, >10E+06 DNA copies per ml of NPS, mRNA or antigen detection in NPS, positive IgM, and low IgG avidity or seroconversion in follow up samples[95].

The diagnosis of acute infections of HBoV2-4 is as complicated as of HBoV1 with viral DNA persistence and serologic cross-reactivity. No clear association of HBoV2 and HBoV3 to GE or other diseases have been found, and the viral DNAs have been detected in stools of both patients and healthy individuals (recently reviewed in [95]).

Competition assay for clarifying cross-reactive IgG EIA results. In HBoV diagnostics the serological cross-reactions of the four HBoVs greatly interfere with the result interpretation and a competition assay has been successfully employed to compensate for the cross-reactivities. Kantola et al. even stated that “Correction for cross-reactivity is a prerequisite for VLP-based HBoV seroepidemiology”[105]. In the competition assay homologous or heterologous antigens are incubated with the sample prior to the EIA, and only the specific antibodies not reacting with the common epitopes of the soluble antigen remain to be detected in the EIA. The result is then compared with the non-competed result. For example, when a sample is analyzed for HBoV1 IgG, the sample is analyzed in two adjacent wells with HBoV1 antigen bound to the well. In the first well, no competition is used, and in the second, the sample is incubated with HBoV2-4 antigens. The HBoV1 IgG-positive sample would give a positive result in both of the wells, while HBoV2-4 IgG positive samples would be negative in the second well due to the soluble HBoV2-4 antigens removing the antibodies.

1.3 HUMAN PROTOPARVOVIRUSES

1.3.1 HUMAN BUFAVIRUS

Discovery of BuVs. In 2012 Phan and colleagues studied fecal samples from under 5-year-old children with acute diarrhea in Burkina Faso in order to find novel enteric viruses[34]. Rotavirus antigen-negative samples were subjected to viral-particle enrichment and NGS, and sequences showing significant similarities to parvoviral proteins were obtained from one sample. After filling the gaps between NGS-derived sequences by PCR and the ends by RACE (Rapid Amplification of cDNA Ends), a nearly full-length genome with complete open reading frames of NS1 and VP1 and VP2 of a novel parvovirus was acquired (4921 nts, GenBank #JQ918261)[34]. The virus was named bufavirus 1-BF96 based on the discovery place and sample code: Burkina Faso parvovirus type 1 strain 96.

In a subsequent PCR study, three more samples harbored BuV DNA giving the overall prevalence of 4% (4/98)[34]. However, when the complete protein-coding regions of all four strains were determined, one showed only 72% amino acid identity on the VP2 protein compared to the other three, although the NS1 was >95% similar. Thus, a second BuV type, BuV2, was proposed[34]. Two additional cohorts of fecal samples from children were also analyzed with the same BuV nested PCR, and 1 of 63 (1.6%) in Tunisia (children with nonpolio acute flaccid paralysis), but none of 100 in Chile (children with diarrhea), harbored BuV DNA. This result showed that the novel virus was not geographically restricted to Burkina Faso[34].

Two years later in 2014, a third genotype, BuV3, was discovered in Bhutan in fecal samples of diarrheic children with a prevalence of 0.8% (3/393)[129]. The BuV3 was nearly identical to BuV1 and BuV2 in NS1 region with >95% similarity at both the nt and aa levels, however, the VP region was more different with only ~73% similarity to the VP2 of BuV1 and 65% to that of BuV2 (aa level). The VP2 of BuV3 had also 3 amino acids more than BuV1 and BuV2[129].

Genome and structure of BuVs. The genome of BuV has two major ORFs that encode for the non-structural NS1 and the capsid proteins VP1 and VP2 (Fig. 6), with the lengths of 672-674, 708-711, and 570-573 amino acids, respectively, depending on the genotype[130]. The genome contains also an ATP- or GTP-binding Walker loop, two

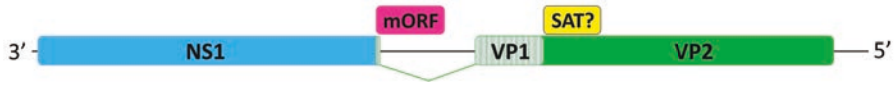


Figure 6. **Organization of the BuV genome.** The illustration is based on BuV1 strain BF96 (GenBank no. JQ918261, 4912 nts). Amino acid alignment of the major NS and VP proteins, the putative short proteins, middle ORF and SAT, as well as the theoretical splicing of VP1 are shown[34,36].

conserved replication initiator motifs, PLA2 in the VP1 unique region, a glycine-rich sequence in the VP2 N-terminal end, and the same splice sites. The three BuVs are nearly identical in the NS1 (94-96% on aa), whereas the capsid is more divergent (VP1: 71-78% on aa; VP2; 65-73% on aa). Besides the major ORFs, a third ORF has been identified in the middle of the genome, and it codes for a smaller ~130 aa-long protein (Fig. 6). This mORF does not share any similarity to other parvoviral proteins, however, the later discovered animal BuVs and human CuV have similar short ORFs in their genomes[131]. The ORF coding for small alternatively-translated protein (SAT), which is found in other protoparvoviruses and amdoparvoviruses, was not originally reported for any of the BuVs, however, Phan et al. describe such an ORF of 82-103 aa in the article reporting the discovery of CuV[36]. Overall, all ORF and mRNA estimations have been made in silico and await confirmation in cell culture experiments.

The capsid structures of the three BuVs have been determined by cryo-electron microscopy, and the capsid surface morphology of the three BuVs were nearly identical (Fig. 7). The 3D structure also resembled closely the morphology of the other parvoviruses in the *Parvovirinae* subfamily and especially those in the genus *Protoparvovirus* (Figs. 4 and 8)[132]. The sequence similarity between BuV1 and the other protoparvoviruses was only ~30%, while the structural similarity was ~75%. The biggest structural difference was located in the 3-fold protrusions: an 11-aa deletion in the variable loop VIII of the BuVs caused a formation of three separate protrusions, whereas in the classical protoparvoviruses the 3-fold protrusions form a single pinwheel structure (Figs. 4, 7 and 8).

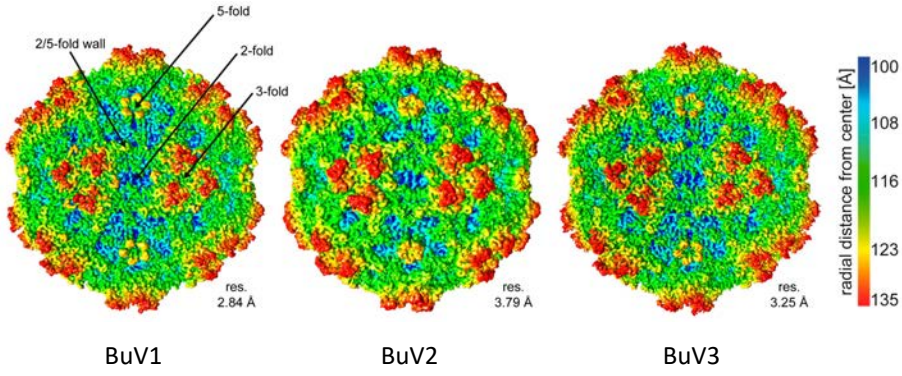


Figure 7. **The capsid structures of BuV1, BuV2 and BuV3.** The capsids are viewed from the 2-fold axes and colored according to radial distance as indicated by the scale bar (from blue [inside] to red [outmost parts]). The three separate protrusions are visible in the 3-fold axes. Res., resolution. Reproduced and modified from Ilyas et al. 2018[132] with permission.

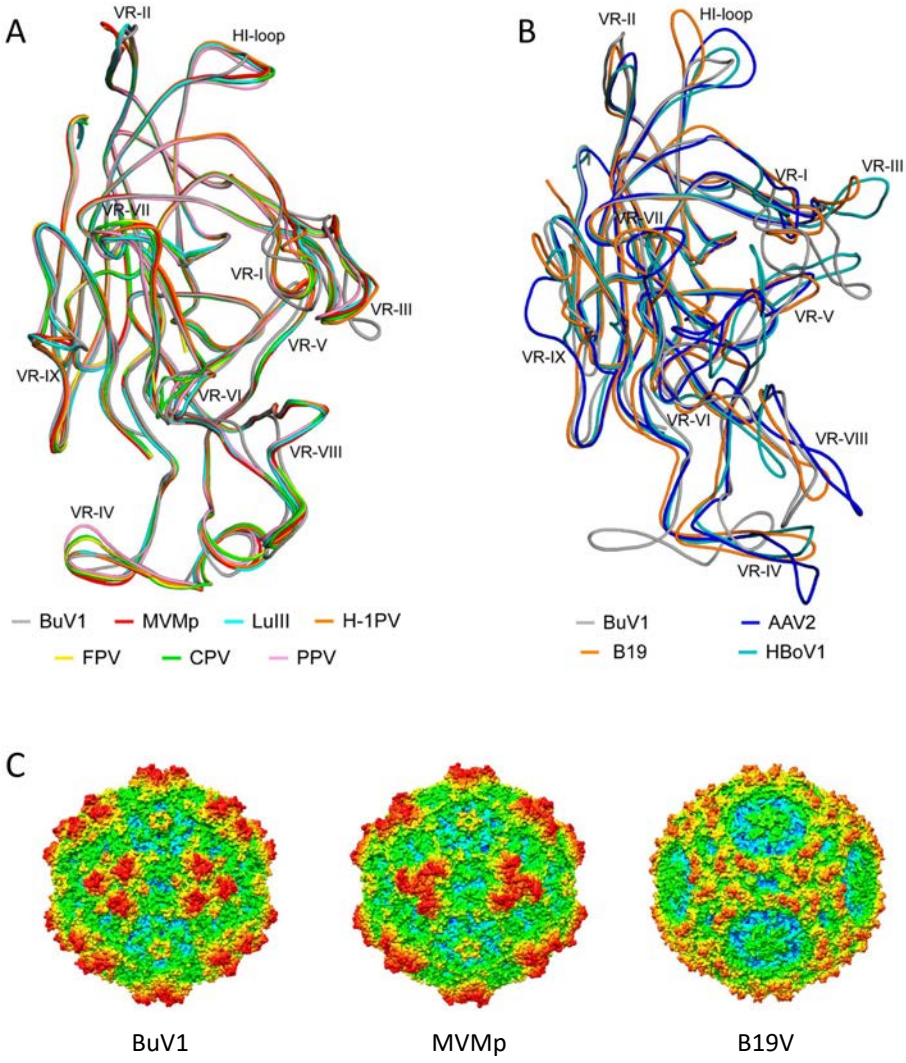
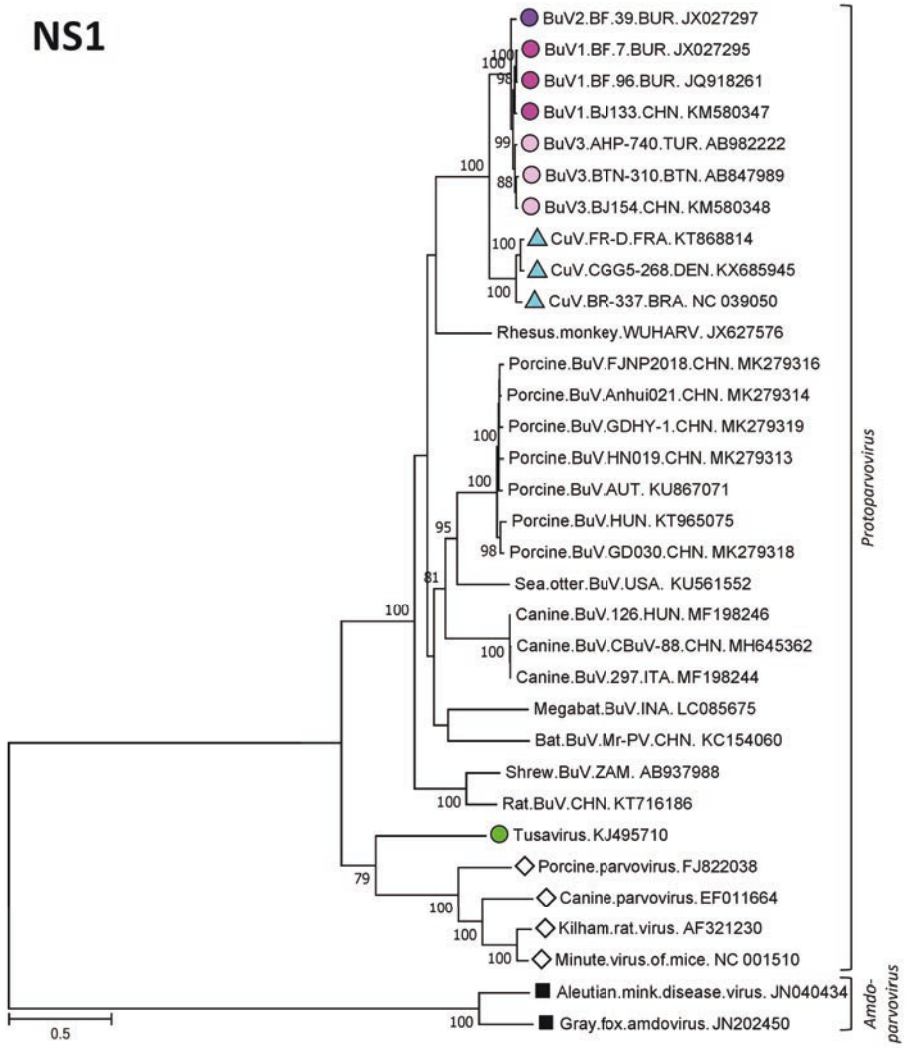


Figure 8. **The structures of BuV1 and selected parvoviruses**. Stereo projection of the structural superposition of VP monomers from (A) BuV1 and other protoparvoviruses, and (B) BuV1 and the human-infecting parvoviruses AAV, B19V and HBoV1. The VR regions (VR-I to VR-IX) and the HI-loop are labeled. The 11-aa deletion that causes the 3-fold pinwheel structural difference between the BuVs and the other protoparvoviruses can be seen in the VR-VIII loop. (C) Radially-colored capsid structures of BuV1, MVMp and B19V, coloring from blue to red as shown in Figure 7. AAV, adeno-associated virus; B19, human parvovirus B19; BuV, human bufavirus; CPV, canine parvovirus; FPV, feline panleukopenia virus; H-1PV, rat protoparvovirus H-1; HBoV, human bocavirus; HI-loop, loop between βH and βI strands (see Fig. 3 for more info); LuIII, parvovirus LuIII; MVMp, minute virus of mice; PPV, porcine parvovirus; VR, variable region. Reproduced and modified from Ilyas et al. 2018[132] with permission.

Phylogenetic analysis. Phylogenetically the three BuV genotypes belong to one species, *Primate protoparvovirus 1* as they are nearly identical in the NS1 protein sequence. In the NS1 region BuVs are closest animal bufaviruses and the classical animal protoparvoviruses that infect rodents, canines and porcines. However, the capsid proteins of the three BuV genotypes are only 65-73% similar (VP2 aa)[133], and in the phylogenetic tree, they cluster in three different branches. In addition, the members of genus *Amdoparvovirus* cluster between BuVs and the classical protoparvoviruses in the VP region, and the *Protoparvovirus* genus seems to be divided (Fig. 9).



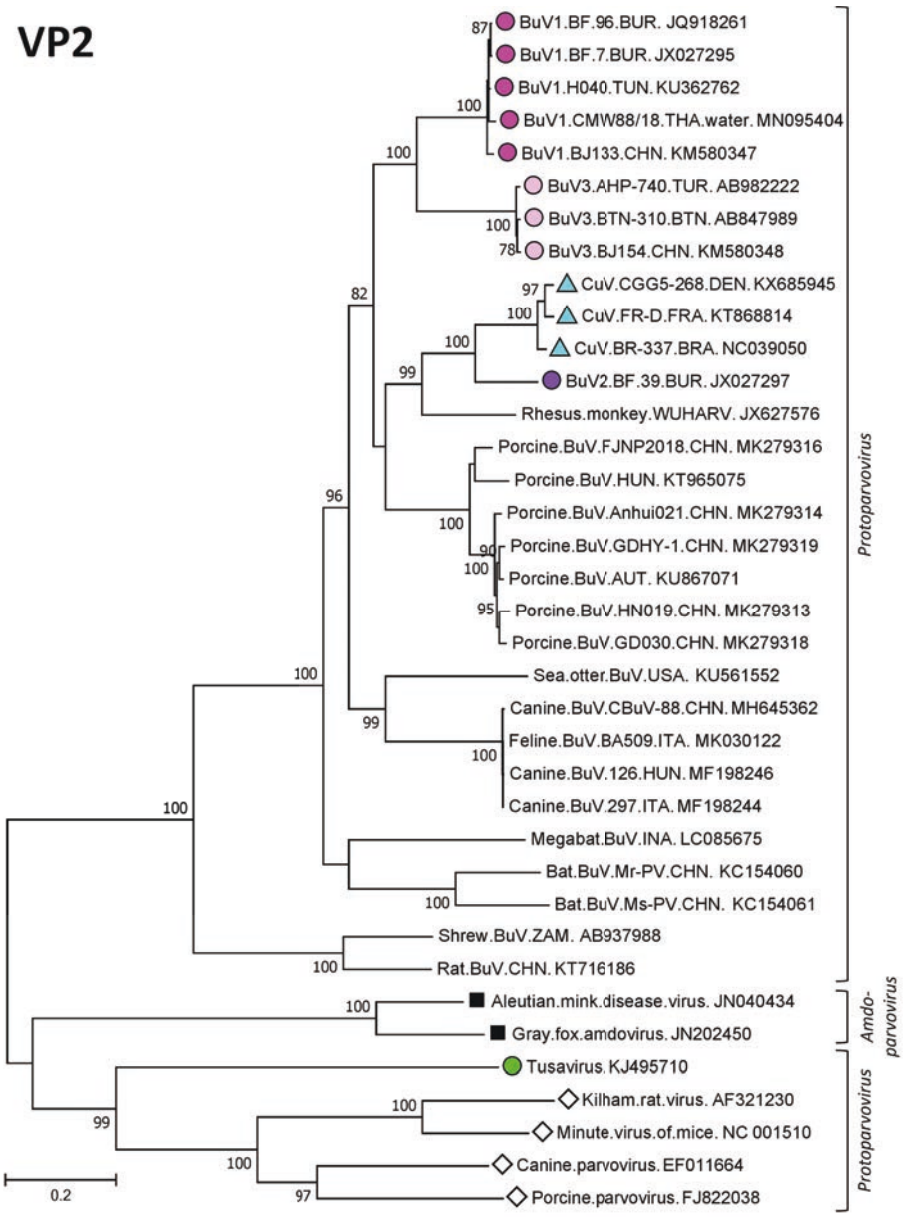


Figure 9. **Phylogenetic trees of NS1 and VP2 of selected human and animal protoparvoviruses, and amdoparvoviruses.** Magenta circles denote BuV1, purple circles BuV2, light magenta circles BuV3, turquoise triangles CuV, green circle TuV, empty diamonds the classical protoparvoviruses, and black squares amdoparvoviruses. The phylogenetic tree was based on nt sequences, and the trees were generated with the maximum likelihood method and the general time reversible model with 1000 replicates using MEGA 7.0. The bootstrap values are shown if >75%[134].

BuV in human samples. To date, BuV DNA has been detected in fecal samples in Africa, Europe, Asia, North America and South America with low prevalences of 0.3-8% in both children and adults, as well as in surface water in Ecuador and Thailand (see Table 3 in Results and discussion). Overall, the findings during the first five years after the BuV discovery (2012-2017) suggested that BuV might have a role in gastroenteritis as all BuV DNA-positive patients had GE symptoms and the non-diarrheal controls were BuV-DNA negative. On the other hand, BuV DNA was not frequently detected in any cohort, the viral loads were low when reported, and the BuV DNA-positive samples often harbored a known GE-causing pathogen. Since 2018, BuV DNA has been occasionally detected in stools of healthy or non-diarrheal individuals[135–137], but large-scale analysis of other sample types than feces is still lacking. In addition, no patient with acute BuV infection has been described in the literature and identifying such patients in the future will be critical for defining the clinical picture of BuV infections. With B19V and HBoVs, it is well established that the viral DNA persists for a long time after primary infection and that the DNA of the respiratory HBoV1 can be detected also in fecal samples. Whether the same phenomenon occurs also with BuVs, is not known.

Interestingly, the DNA of BuV1 and/or BuV3 has been detected in several studies around the world, but DNA of BuV2 has been absent. The BuV2 DNA-positive sample from first study identifying BuV in Burkina Faso is still the only sample with BuV2 DNA described in the literature. However, feces, the sample type used in most studies, may not be a good sample for detecting BuV2 DNA. In addition, there might be other issues with BuV2 detection, for example assay specificity and sensitivity for BuV2 versus BuV1 and BuV3 that hamper the detection.

1.3.2 HUMAN TUSAVIRUS

In 2014, a second novel protoparvovirus, tusavirus (TuV), was detected in human stool[35]. Phan and coworkers analyzed fecal samples with metagenomic deep sequencing combined with viral enrichment from infants and children with unexplained diarrhea in Tunisia to find novel viruses potentially contributing to the illness. Although the 180 samples had been tested negative for rotavirus, norovirus, astrovirus, sapovirus, adenovirus types 40 and 41, and Aichi virus by RT-PCR, many reads matching these viruses were identified by deep sequencing, of which sapovirus was the most frequent[35].

One pool showed a single read with similarity to the NS1 of rat parvovirus. The individual sample containing the viral read was identified by PCR, and by further deep sequencing and PCR, a nearly complete genome of a novel parvovirus (4 424 nt) was acquired (Fig. 10). The virus was named tusavirus (Tunisian stool-associated parvovirus). No other eukaryotic viral sequences were found in the NGS reads of the sample that could explain the symptoms of the child, however, no testing optimized for pathogenic bacteria or parasites was carried out. Nested PCR analysis of the other samples remained TuV negative, and the overall prevalence of TuV in Tunisian children was 0.56% (1/180)[35].

Figure 10. Organization of the TuV genome. The illustration is based on strain Tu491 (GenBank no. KJ495710, 4424 nts). Amino acid alignment of the major NS and VP proteins, as well as the theoretical splicing of VP1 are shown[35]

Phylogenetically the Kilham rat parvovirus shares the highest identity to TuV (44% in NS1), and TuV clusters together with the classical animal protoparvoviruses, such as the mouse pathogen MVM and the canine pathogen CPV, instead of BuVs (Fig. 9). According to the genus and species demarcation criteria of the ICTV, TuV was proposed to belong to a novel species in the *Protoparvovirus* genus[2,35], however the classification is still pending (in January 2020)[2].

Since the discovery, TuV-like sequences have been described only in one article: partial sequences of 438-1519 bps matching 39-82% to TuV sequence were obtained by NGS in feces from fur seals in Brazil (see also Table 1, animal bufaviruses)[138]. Otherwise, the studies of TuV are limited to the ones included in this thesis, and it is still unknown whether TuV is a human virus.

1.3.3 HUMAN CUTAVIRUS

Discovery of CuV. Human cutavirus was discovered in 2016 by an identical process to that of the BuV and TuV: CuV sequence reads were identified with metagenomics in diarrheal fecal samples from children in Brazil[36]. Further prevalence studies were carried out with PCR, which revealed CuV DNA in fecal samples of 4/245 (1.6%) children in Brazil as well as in 1/100 (1%) children in Botswana.

Then the story of CuV changed. During previous years several metagenomics datasets had been generated from human, animal and environmental samples, and these data sets were screened for the novel CuV in silico[36]. Unexpectedly, one pool of skin biopsies from cutaneous T-cell lymphoma (CTCL) (n=2) and parapsoriasis (n=2) patients from France contained nearly identical sequences to CuV. PCR analysis of individual samples in the pool revealed CuV DNA in both of the CTCL tissues, while the parapsoriasis samples were negative. Both CTCL patients had mycosis fungoides (MF), the most common subtype of CTCL. In an additional PCR study, two more MF tissues were shown to harbor CuV DNA, whereas all other skin biopsies including ten skin carcinoma, eight parapsoriasis, eight eczema or eczematoid dermatitis and three healthy skin samples were negative for CuV DNA[36]. The presence of CuV in the MF biopsies was further confirmed with in situ hybridization showing CuV in 2/4 CTCL samples initially positive for CuV DNA. The ISH showed only a few, but highly positive cells harboring CuV[36]. Due to the link to CTCL, the virus was finally named cutavirus (cutaneous T-cell lymphoma associated parvovirus).

Genome and structure of CuV. The genome of CuV has two major ORFs coding for NS1 and VP1/2 as well as a shorter middle ORF and a SAT-like ORF (Fig. 11)[36,139]. BuV2 is phylogenetically the closest virus, sharing 76% similarity in NS1 and 82% on VP1. The mORF of CuV is shorter than in BuVs and has only 45% similarity. The VP1u contains PLA_2 activity, and interestingly, the N-terminal end of VP2 shows the highest divergence between the CuV strains (Fig. 12). The CuV capsid structure has not been determined, however, due to the close sequence resemblance of CuV and BuV2 VP2s, the structure could be hypothesized to be very similar[132].

Figure 11. **Organization of the CuV genome**. The illustration is based on strain BR-337 (GenBank no. NC_039050, 4456 nts). Amino acid alignment of the major NS and VP proteins, the putative short proteins (middle ORF and SAT), as well as the theoretical splicing of VP1 are shown[36]. VP2 and SAT are translated in different frames.

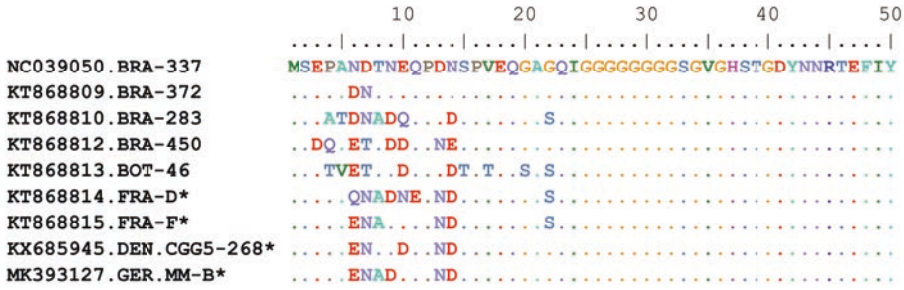


Figure 12. **Alignment of the first 50 amino acids in the N-terminal end of CuV VP2 protein.** The alignment contains all published CuV strains on January 2020. Although the entire VP2 is nearly identical between the CuV types (97% similarity on aa), most of the differences are located before the glycine-rich region. The first M in the alignment is the starting methionine of VP2. *, CuV from cancer biopsy; dot (.), identical aa compared to the upper sequence.

Phylogenetic analysis. Nearly full-length genomes with complete NS and VP ORFs have been amplified from four patients: in a fecal sample of a Brazilian child, in CTCL biopsies of two French CTCL/MF patients and in one Danish melanoma patient[36,139]. Partial genomes, missing only the N-terminal end of NS1, were obtained from four other children (three from Brazil, one from Botswana)[36]. The different strains have been very similar to each other with 93-98% aa identity in NS1 and 96-99% in VP1, and it seems that there is only one genotype of CuV so far[140]. Phylogenetic analysis showed that CuV was closely related to human BuVs, however as the NS1 was only 76% similar on aa to the closest human BuV, BuV2, a distinct protoparvovirus species was suggested. Currently

CuV is classified to species *Primate protoparvovirus 3* in genus *Protoparvovirus* by ICTV parvovirus study group (as in January 2020)[2].

Interestingly, although based on NS1, BuV1-3 and CuV are clearly different viruses, the capsid protein region of CuV clusters with that of BuV2, while the VPs of BuV1 and BuV3 cluster together. Furthermore, when comparing the phylogenetic trees of the entire *Protoparvovirus* genus with that of the closest genus *Amdoparvovirus*, in the NS1 tree the two genera are separate, whereas in the VP1 or VP2 trees amdoviruses are between the classical protoparvoviruses and bufaviruses (including human and animal BuVs, and CuV) (Fig. 9).

CuV in human samples. The finding of CuV DNA in cancerous tissue directed the forthcoming studies towards skin biopsies from different patient groups instead of fecal samples, as was the case with human BuVs. Only one additional study describing CuV in fecal samples has been published: Siqueira et al. identified NGS reads matching CuV in 2 children from two isolated Amazonian villages in Venezuela[135].

Currently CuV DNA has been detected in melanoma tumor biopsies in Denmark (1/10, 10%) and in Germany (2/179, 1.1%), in MF lesions in Germany (6/71, 8.5%) as well as in skin swabs of healthy (9/237, 3.8%) and of HIV-positive men (35/205, 17.1%)[139,141,142]. In one study from Italy, no CuV DNA was detected in CTCL skin samples (n=55), however, the methods and value of this study published in an Italian journal are difficult to interpret[143]. Although Kreuter et al. detected CuV DNA only in MF lesions, and all other studied cutaneous T-cell, NK-cell and B-cell lymphomas (n=59) were negative, more studies are needed to determine to role of CuV in CTCL. In general, parvoviruses are not oncogenic like some other viruses, but rather oncolytic such as the rat protoparvovirus H-1 (H-1PV)[144,145], which has been studies as a cancer therapeutic. Or parvoviruses might be oncotropic as they need rapidly dividing cells for replication. In addition, B19V has been shown to remain in skin tissue years or even decades after primary infection without apparent harm or symptoms to the individual. CuV could show similar characteristics. However, for a pre-disposed individual, such as immunocompromised or cancer patients, a persisting virus might cause a disease.

1.4 ANIMAL BUFAVIRUSES

In October 2012, two months after the initial article describing human BuV, the first animal BuV was identified[146]. Handley and coworkers studied the effect of the simian immunodeficiency virus (SIV) infection to the enteric virome, enteropathy, and AIDS progression in rhesus monkeys, and discovered a novel parvovirus resembling BuV2 (Fig. 9). The novel virus was named WUHARV parvovirus. By NGS and PCR analysis, viral DNA was detected in the feces of ten SIV-infected rhesus monkeys (10/35, 28.6%) and in one non-infected rhesus monkey (1/51, 1.9%)[146]. In addition, the same virus was present

in the blood of four SIV-infected animals euthanized due to advanced AIDS[146]. Whether the viremia was a result of a true systemic infection or that the virus had invaded the enteric tissues and leaked into the blood stream as the immune system of the animals were severely harmed by AIDS, cannot been concluded from this study. Further studies are also needed to pinpoint the role of the virus in healthy rhesus monkeys. A control group of 54 african green monkeys (29 SIV+, 25 SIV-) were all negative for the new virus.

During the following years several animal BuV-like viruses have identified in various animal species around the world, both in wild and domestic animals. BuV-like viruses were described in shrews and baboons in Zambia[147], in bats in Hungary, China and Indonesia[148–150], in rats in China[133], in domestic pigs in Austria, Hungary, and China[151–153], in sea otters in the USA[154], and the most recently, canine BuV was detected in domestic dogs in Italy, Hungary and China as well as in domestic cats, wild wolves and wild foxes in Italy[155–159]. In most cases these viruses have been searched for in fecal samples or in rectal, nasal or oropharyngeal swabs, however, animal BuVs have been detected also in the spleen of shrews, baboons, megabats, and sea otters, and in serum of dogs pointing to systemic infections in these animals[150,154,156,157]. Phylogenetically the animal bufaviruses cluster together with human BuVs and CuV, both in the NS as well as in the VP regions.

The causative role of BuVs in animal disease and possible symptoms in the infected animals warrants further studies. In many of the published studies, animals have been first screened for novel viruses by NGS, and then a prevalence study has been carried out by PCR including the same animal cohort and occasionally some other animals. Although the approach gives a good first impression of the prevalence of the virus and perhaps directs future studies, the etiological role of these viruses in any disease cannot be reliably assessed. In one study of dogs in Italy some correlation of canine BuV and a disease was obtained: the presence of canine BuV DNA in the nasal and oropharyngeal swabs was clearly associated with canine infectious respiratory disease (kennel cough)[155]. Overall, the detection rates of the DNA of BuV-like viruses in the corresponding animals are higher compared to BuV in human samples (Table 1).

Table 1. ***Studies describing BuV-like viruses in diverse animal species.*** *The table presents all studies (Jan 2020) and is in chronological order*

Study	Virus	Animal species	Sample type	Country	Sampling time	n	Positive (%)	Detection method	Symptoms, other info
Handley 2012[146]	WUHARV parvovirus	**Rhesus monkeys**	Feces	USA	Unknown	35 SIV+	**10 (28.6%)**	NGS, PCR	All NGS pos detected also by PCR
		Rhesus monkeys	Feces	USA	Unknown	51 SIV-	**1 (1.9%)**	NGS, PCR	NGS pos, PCR neg
		Rhesus monkeys	Serum	USA	Unknown	10 SIV+, euthanized	**4 (40%)**	PCR	Euthanized due to advanced AIDS
		African green monkey	Feces	USA	Unknown	29 SIV+; 25 SIV-	**0**	NGS	Afr. green monkey develops persistent high-level viremia but not AIDS
Kemenesi 2015[148]	BtBV/V3/HUN/2013 BtBV/V7/HUN/2013	**Bats (*M. schreibersii*)**	Feces	Hungary	2013	Unknown	**Unknown**	NGS	26 reads close to parvoviruses
						13	**5 (38.5%)**	PCR	
Sasaki 2015[147]	MpBuV (Mpulungu BuV, later changed to Eulipotyphla protoparvovirus 1)	**Shrew (*Crocidura hirta*)**	Intestinal content	Zambia	2012	23	**5 (21.7%)**	PCR for MpBuV	PCR-positive shrews -> lung, spleen, liver, kidney tested separately by PCR
			Spleen			*5*	***5 (100%)***	"	*Tissue from the 5 MpBuV positive animals analyzed separately*
			Liver			*5*	***4 (80%)***	"	
			Lung			*5*	***0***	"	
			Kidney			*5*	***0***	"	
		Baboon	Spleen	Zambia	2009 - 2011	100	**3 (3%)**	Deg nPCR	
		Vervet monkey	Spleen	Zambia	2009 - 2011	89	**0**	Deg nPCR	
		Shrew	Spleen	Zambia	2011 - 2013	49	**17 (34.7%)**	Deg nPCR	Includes the 5 above. Two shrew species: *C. hirta* and *C. luna*
		Rodents	Spleen	Zambia	2011 - 2013	298	**0**	Deg nPCR	
Yang 2016[133]	RatBuV	**Wild rats**	Intestinal content	China	Jun - Aug 2004	40, 4 pools	**Unknown**	NGS	61 reads close to WUHARV
		Wild rats	Intestinal content	China	Jun - Aug 2004	40	**5 (12.5%)**	nPCR	
Wu 2016[149]	4 new bat BuVs	**Bats, 40 representative species**	Pharyngeal and anal swabs	China, across 29 provinces	Oct 2010 - Oct 2013	4440 bats, 84 pools	Unknown	NGS, PCR verification	4 new bat BuVs clustering with human BuVs and WUHARV

Study	Virus	Animal species	Sample type	Country	Sampling time	n	Positive (%)	Detection method	Symptoms, other info
Hargitai 2016[151]	Swine/Zsana3/2013/ HUN	**Domestic pigs**	Feces	Hungary	Jan - Sep 2013	21	**19 (90.5%)**	NGS, PCR	4 reads of BuV by NGS. PCR-screening of healthy sows and piglets with posterior paralysis
		Domestic pigs	Feces	Hungary	Nov 2008	15	**5 (33.3%)**	PCR	Healthy piglets from control farm (4-weeks-old all neg, 3-month-old all pos)
Kluge 2016[138]	BuV-like sequences (and TuV-like sequences)	**Fur seal, South American**	Feces*	Brazil	2012 - 2013	5, pooled to one	Unknown	NGS	1 contig: 460nt, 43% on NS1 of Solwesi shrew BuV. *TuV: 2 contigs: 616nt 46% on VP2 and 344nt 82% on NS1 of TuV*
		Fur seal, Subantarctic	Feces*	South Brazil cost	2012 - 2013	5, pooled to one	Unknown	NGS	3 contigs: 1.6kb 57% and 319nt 41% on NS1 of BatBuV CHN; 322nt 36% on VP2 of MgBuV. *TuV: 3 contigs: 1.5kb 39%, 612nt 39%, 438nt 42% on VP of TuV*
Sasaki 2016[150]	Megabat BuV1 Megabat BuV2 Megabat BuV3	**Megabat**	Spleen	Indonesia	2010 - 2014	183	**4 (2.2%)**	Deg nPCR	One bat positive in both spleen and feces
		Megabat	Feces	Indonesia	2011 - 2014	96	**9 (9.4%)**	Deg nPCR	
Liu 2016[152]	Porcine BuV (strain 61)	**Domestic pigs**	Feces	Austria, Hungary	Jan - Sep 2012	6 sample pairs	**Unknown**	NGS	6 farms, 1 sample from healthy pig, 1 with diarrhea per farm. Several contigs from two farms (1 AUT, 1 HUN)
		Domestic pigs	Feces	Austria, Hungary	Jan - Sep 2013	60	**8 (13.3%)**	real-time PCR	All 8 positives from 2 farms in AUT (NGS pos HUN farm was PCR neg)
Siqueira 2017[154]	Sea otter parvovirus	**Sea otter**	Liver, Lung, lymph node, spleen	California, USA	2000 - 2013	69 animals	**42 (60.9%)**	NGS and nPCR	Found dead or died under care. 42 animals pos -> 78% pos in mesenteric lymph node (29 pos of 37 available lymph nodes), other organs <10%
Martella 2018[155]	Canine BuV	**Domestic dog**	NOP	Italy	2011 - 2015	58	**18 (31%)**	qPCR	<1 y old, canine infectious respiratory disease
			NOP	Italy	2012 - 2015	90	**0**	qPCR	<1 y old, healthy
			Enteric[#]	Italy	2010 - 2015	81	**26 (32.1%)**	qPCR	<1 y old, gastroenteritis
			Enteric[#]	Italy	2011 - 2015	78	**15 (19.2%)**	qPCR	<1 y old, healthy
			Enteric[#]	Hungary	2012	20	**8 (40%)**	qPCR	<1 y old, gastroenteritis
			Enteric[#]	Hungary	2012	40	**19 (47.5%)**	qPCR	<1 y old, healthy

Study	Virus	Animal species	Sample type	Country	Sampling time	n	Positive (%)	Detection method	Symptoms, other info
Diakoudi 2019[158]	Feline BuV (sequence 99.9% identical to canine BuV)	**Cat**	NOP	Italy	2016 - 2017	180	**22 (12.2%)**	qPCR	both with and without respiratory signs
			Enteric[#]		2012 - 2015	90	**2 (2.2%)**	qPCR	gastroenteritis
			NOP		2012 - 2013	179	**13 (7.3%)**	qPCR	respiratory signs
			NOP		2012 - 2013	125	**16 (12.8%)**	qPCR	no respiratory signs
Li 2019[156]	Canine BuV	**Domestic dog**	Feces	China	Feb 2016 - Jun 2018	121	**51 (42.15%)**	nPCR	diarrheal dogs
			Plasma		Jan - Aug 2017	11	**1 (9.1%)**	nPCR	diarrheal dogs
			Feces		Jun 2018	15	**0**	nPCR	healthy dogs in one kennel
			Plasma		Jan - Aug 2017	5	**0**	nPCR	healthy
Sun 2019[157]	Canine BuV	**Domestic dog**	Serum	China	2016 - 2018	200	**5 (2.5%)**	nPCR	CPV PCR pos
			Serum			200	**0**	nPCR	CPV PCR neg
			Diarrheal feces			80	**5 (6.25%)**	nPCR	canine distemper n=39, and CPV n=41 (all canine BuV pos also CPV pos)
			Nasal swab			60	**0**	nPCR	unknown disease
Melegari 2019[159]	Canine BuV	**Wolf**	Rectal swab	Italy	Sep 2009 - Apr 2017	41	**7 (17.1%)**	qPCR	necropsied wolves found dead after traumatic events
		Fox	Rectal swab		Sep 2009 - Apr 2017	86	**9 (10.5%)**	qPCR	red foxes shot during regular hunting seasons
Sun 2020[153]	Porcine BuV	**Domestic pigs**	Serum	China	Dec 2017 - NoV 2018	292	**41 (14.0%)**	nPCR	Healthy pigs from 112 commercial farms in six provinces. Between the provinces, the porcine BuV DNA prevalence ranged from 10.7% to 23.0%
			Feces	China	Dec 2017 - NoV 2018	92	**23 (25.0%)**	nPCR	

*, Feces collected from intestines of deceased animals; #, Feces and rectal swabs

AUT, Austria; CHN, China; CPV, canine parvovirus; deg nPCR, degenerative nested PCR; HUN, Hungary; neg, negative; nPCR, nested PCR; NOP, nasal and oropharyngeal; qPCR, quantitative PCR; pos, positive; SIV, simian immunodeficiency virus; y, year

2 AIMS OF THE STUDY

This thesis project started when the first human protoparvovirus, BuV, was discovered in 2012. The overall aim was to study the epidemiology and clinical relevance of this novel virus as well as to provide evidence of whether or not the virus is truly a human virus. The first specific aim of this thesis was

1. To set up a quantitative PCR (qPCR) method for analyzing BuV DNA in human samples to elucidate the BuV occurrence in Finland

During Study I, the third genotype of BuV, BuV3, as well as another putative human protoparvovirus, TuV, were described. These viruses were included in the studies, and the next aims were

2. To clone the putative VP2 genes of BuV1, BuV2, BuV3 and TuV and generate virus-like particles (VLPs) to be used in IgG EIA assays
3. To determine the BuV1-3 and TuV IgG seroprevalence in Finland, and assess the potential cross-reactivity of these closely related viruses

In 2016, a third putative human protoparvovirus, CuV, was discovered. Besides in feces, CuV DNA was detected in human skin biopsies, and this prompted the final studies of this thesis

4. To clone the putative VP2 gene of CuV, generate VLPs, and add CuV to the protoparvovirus IgG EIA panel
5. To determine the BuV1-3, TuV and CuV IgG seroprevalence worldwide
6. To assess the CuV IgG cross-reactivity with BuV1-3 and TuV
7. To set up qPCR methods for TuV and CuV DNA detection
8. To search for BuV, TuV and CuV DNA in skin biopsy samples from cutaneous T-cell lymphoma (CTCL) and transplant patients and from healthy adults in Finland as well as determine the IgG seroprevalence of all three viruses in these patient cohorts

3 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 STUDY COHORTS AND CLINICAL SAMPLES

Cohort 1, fecal samples from GE patients in Finland (I)
The GE cohort consisted of fecal samples from 629 gastroenteritis patients of all ages (median 51.5 years, range 0–99) in Finland. The samples were obtained retrospectively from routine diagnostics for bacterial (n=243) and viral (n=386) gastroenteritis-inducing pathogens (HUSLAB, Helsinki University Hospital, Helsinki, Finland). The samples were sent to HUSLAB from diverse locations within Finland and were collected either during October 2012–March 2013 (bacterial samples) or in April–May 2013 (viral samples).

Cohort 2, fecal, nasal swab and serum samples from children in Finland (II)
The Tampere pediatric cohort consisted of 955 patients with symptoms of acute gastroenteritis (AGE, n = 172), acute respiratory tract infection (ARTI, n = 545) or both (n = 238). The recruited children (median 14 months, range 6 days – 15.6 years) were either visiting an outpatient clinic (n=137, 14.3%) or were admitted to a hospital ward (n=818, 85.7%) between September 2009 and August 2011 in Tampere, Finland. Fecal and nasal swab samples were obtained from all patients and serum samples from 228 of whom blood sampling was medically indicated.

Cohort 3, serum or plasma samples from healthy adults in Finland (II)
The student/staff cohort consisted of serum or plasma samples from 180 healthy adults (median age 30 years, range 19–65) collected between 2007 and 2013 in Helsinki, Finland. All 84 medical students from the University of Helsinki and 79/96 staff members from the Helsinki University Hospital or from the Faculty of Medicine, University of Helsinki, were of Finnish descent, the remaining 17 staff members were originally from Asia (12), the Americas (1) or other European countries (4). Follow-up samples from BuV IgG-positive individuals were obtained (frozen-stored) from our archive or collected from participants.

Cohort 4, serum samples from veterinarians in Finland (III)
The veterinary cohort consisted of serum samples from 324 healthy adults (median age 40.2 years, range 19–79) collected from participants at the national Annual Veterinary Congress in 2009 in Helsinki, Finland. The cohort consisted of veterinarians (82%), veterinary students, and veterinary nurses, and most (92%) completed an electronic questionnaire of background information.

Cohort 5, serum samples from the United States (III)
The USA cohort consisted of serum samples from 84 healthy blood donors (median age 41.3 years, range 18–72) obtained from the Blood Systems Research Institute (San Francisco, CA, USA). The samples were collected during April 2009 in 2 locations: Arizona (n = 40) and Mississippi (n = 44).

Cohort 6, serum samples from Iraq (III)
The Iraq cohort consisted of serum samples from 99 healthy adults (median age 39.7 years, range 18–60) collected in November-December 2013 in Nasiriyah, Dhi Qar, in southern Iraq.

Cohort 7, serum samples from Iran (III)
The Iran cohort consisted of serum samples from 107 healthy adults (median age 42.2 years, range 18–77) obtained from the Hamadan Blood Transfusion Organization, Hamadan, Iran. The samples were collected during 2015-2016.

Cohort 8, serum samples from Kenya (III)
The Kenya cohort consisted of serum samples from 107 children (median age 6.9 years, range 0.5-17.8) and 119 adults (median age 43.3 years, range 18–88) who had a febrile illness of unknown cause and had visited health clinics in Wundanyi, Mwatate, or Voi in Taita Taveta County in southern Kenya. All patients completed a questionnaire to provide information of their backgrounds and current symptoms.

Cohort 9, tissue and serum samples from CTCL patients (IV)
The CTCL cohort consisted 25 patients (median age 64 years, range 32–89 years), of whom 23 had mycosis fungoides and 2 leukemic Sezary syndrome. A FFPE skin biopsy from CTCL lesion site was available from all. Additional FFPE tissue biopsies were obtained from the pathology archives from 2 patients with CuV DNA in the original CTCL biopsy: 2 prostate samples were available from one and 12 samples from the other, including prostate, lymph node, melanoma, CTCL and other skin samples (Table 2 in Study IV). Serum samples were available of a separate cohort of 42 CTCL patients, including 3 whose skin biopsies were analyzed.

Cohort 10, tissue and serum samples from transplant patients (IV)
The transplant cohort consisted of 136 immunosuppressed solid-organ transplant patients (median age 62 years, range 22–83 years, data not available from 5 patients), who had received an organ transplant more than 5 years ago (median 10 years, range 5–26 years). Most of the patients (131/136) received liver transplants, 4 received kidney, and 1 heart transplant. A fresh skin biopsy was taken from all: healthy skin biopsies were available from 133 patients and suspected malignant or premalignant skin biopsies from 19 patients (19 carcinoma, 1 melanoma, 4 actinic keratosis; 1–4 biopsies per patient). A

corresponding serum sample was available from 123 patients, and additional 1–3 serum samples from the CuV, BuV, or TuV sero- or genopositive patients were obtained at various time points before the skin biopsy. Overall, serum samples were available from 124 patients.

Cohort 11, tissue and serum samples from healthy adults (IV)
The healthy adult control cohort consisted of 98 immunocompetent adults (median age 48 years, range 18–67 years). A skin biopsy (1-2 per person, 159 in total) were obtained from epicutaneous test areas for allergens or irritants. None had known skin malignancies, although 15 adults had mild skin conditions such as atopic dermatitis. A corresponding serum sample was available from 78 patients.

3.1.2 ETHICAL ASPECTS

All studies were conducted in accordance with the relevant guidelines and regulations. A written informed consent was obtained from participants or from the guardian of the child with cohorts 2-4, 6 and 8-11. The Ethics Committee of the Hospital District of Helsinki and Uusimaa approved the studies using cohorts 1, 3, 4 and 9-11, and the Ethics Committee of Pirkanmaa Hospital District the study using cohort 2. The National Supervisory Authority for Welfare and Health (Valvira) approved studies using cohorts 9-11. Ethics Committees of Medical Sciences at Basrah University and the Al-Hussein Teaching Hospital approved the use of cohort 6, and the Kenyatta National Hospital–University of Nairobi Ethics and Research Committee the use of cohort 8. Informed consent was waived for the deidentified blood donor samples in cohort 7, and the study was approved by the Ethics Committee of Hamadan University of Medical Sciences, Hamadan, Iran. Under US human and health service regulations, the study of preexisting, deidentified samples is not classified as human subject research, and therefore a separate approval for the use on cohort 5 was not applied.

3.1.3 PLASMIDS

For setting up the qPCR and later to serve as a positive control and in creating the standard curve for quantification, a 3.5kb piece of the BuV1 genome was cloned from the original BuV1 DNA-positive fecal supernatant (BF.7) with primers “BuV1-2 NS fwd KpnI” and “BuV1 VP mid rev ApaI” (Table 2). The amplicon was cloned into pSTBlue vector with KpnI and ApaI restriction sites.

For qPCR set up and optimization, a 830-nt piece of the BuV2 genome was cloned from the original BuV2 DNA-positive fecal supernatant (BF.39) with primers “BuV1-2 NS fwd KpnI” and “BuV rev” (Table 2).

For VLP expression, the putative *VP2* genes of BuV1, BuV2, BuV3, TuV and CuV were amplified from the original virus DNA-positive fecal supernatants[34–36,129] with primers

listed in Table 2. The obtained gene was cloned into the Bac-to-Bac system donor vector pFastBacDual (Invitrogen) at the BamHI and SalI restriction sites under the strong *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) polyhedrin promoter. In addition, the *VP2* genes of TuV and CuV were blunt cloned into pSTBlue (Novagen/Merck) to serve as positive controls and for the qPCR set up.

All cloned sequences were confirmed by sequencing.

3.1.4 CELL LINES

Chemically competent cells of *Escherichia coli* -strains DH5α (New England Biolabs) or E. cloni 10G (Lucigen) were used for routine cloning.

Escherichia coli -strain DH10Bac™ (Bac-to-Bac Baculovirus Expression System, Invitrogen), carrying the AcMNPV parent bacmid and a recombination helper vector, was used for creating recombinant baculoviruses carrying the *VP2* genes of BuV1, BuV2, BuV3, TuV and CuV.

Spodoptera frugiperda (Sf9) insect cells were used in creating the recombinant baculoviruses, and *Trichoplusia ni* (HighFive) insect cells were used for VLP expression.

Table 2. ***Primers used in this thesis.*** *Restriction sites are indicated by lowercase letters*

Name	Sequence (5' to 3')	GenBank accession no; oligo position	Use	Study
BuV fwd	ACAGTGTAGACAGTGGATTCAAACTT	JX027295; 705 to 730	qPCR	I, II, IV
BuV rev	GTTGTGGTTGGATTGTGGTTAGTTC	JX027295; 806 to 830	qPCR, plasmid cloning	I, II, IV
BuV NS1 probe	FAM-CGGAAGAGATTTTGACAGTGCYTAGCAA-BHQ1	JX027295; 762 to 789	qPCR	I, II, IV
TuV fwd	CCAGAAAGCCGTATCACCAT	KJ495710; 3085 to 3104	qPCR	IV
TuV rev	AACCAAGTGTTTCTGATCTTATTGCT	KJ495710; 3177 to 3202	qPCR	IV
TuV VP2 probe	TxRd-ACACCAACAATCAACTGCCATACACACC-BHQ2	KJ495710; 3146 to 3173	qPCR	IV
CuV fwd	TAACACATCCCAGAATYGTCACATA	KT868811; 4245 to 4269	qPCR	IV
CuV rev	TTCCATTGTCTTGGAGTGCG	KT868811; 4316 to 4335	qPCR	IV
CuV VP2 probe	JOE-AGTTKTCCTGACCACCAGAAGGTTCCA-BHQ1	KT868811; 4270 to 4296	qPCR	IV
BuV1-2 NS fwd KpnI	ATAggtaccATGGCTCTCAGCAAAGATGTCA	JX027295; 2 to 23	Plasmid cloning, qPCR confirmation	I
BuV NS 1kb rev	TGTTTAAGTTTGCCTGGATGTTC	JX027295; 1058 to 1080	qPCR confirmation	I
BuV1 VP mid rev ApaI	TAgggcccGGTCCATGAGTCTCCTGTTCTTAG	JX027295; 3554 to 3577	Plasmid cloning	I
BuV1 VP2 fwd BamHI	TAggatccATGACTGACACACAAGATGTATCTGA	JX027295; 2786 to 2811	VP2 cloning, qPCR confirmation	I, II
BuV1 VP2 rev SalI	ATTgtcgacTCCATTTTAGATTGTGTAGTTAGGCATAC	JX027295; 4473 to 4495/4501*	VP2 cloning, qPCR confirmation	I, II
BuV2 VP2 fwd BamHI	TAggatccATGTCTGAAAGCAATGAAATTGGAG	JX027297; 2791 to 2815	VP2 cloning, qPCR confirmation	I, II
BuV2 VP2 rev SalI	ATTgtcgacTTACATTGTGTAGTTAGGCATGGCTCT	JX027297; 4474 to 4500	VP2 cloning, qPCR confirmation	I, II
BuV3 VP2 fwd BamHI	TAggatccATGTCCGAAAGCAATGAAATTGACG	AB847989; 2791 to 2815	VP2 cloning	II
BuV3 VP2 rev SalI	ATTgtcgacTTAGTATGTGTAGTTTGGCATTGCTC	AB847989; 4484 to 4509	VP2 cloning	II
TuV VP2 fwd BamHI	TAggatccATGGCAGCCTCTAGCTCAGACAGTG	KJ495710; 2659 to 2683	VP2 cloning	II
TuV VP2 rev SalI	ATTgtcgacTTAGTAAACAGTAGAGGGTACAATTCTTC	KJ495710; 4356 to 4328	VP2 cloning	II
CuV VP2 fwd Br337 BamHI	TAggatccATGTCAGAACCAGCTAATGATAC	KT868811; 2747 to 2769	VP2 cloning	III
CuV VP2 rev Br337 SalI	CTCgtcgacTTACAATGTGTAGTTTGGTAGACA	KT868811; 4433 to 4456	VP2 cloning	III
CuV 583 fwd	TCAACAACTGAAGGMACCAGACTAAC	KT868811; 3753 to 3777	Longer amplicon of CuV VP	IV
CuV 230 fwd	GGAGCCATCTGGGGAAAATAYC	KT868811; 4103 to 4124	Longer amplicon of CuV VP	IV

* the BuV1 VP2 ORF ends at nt 4495, however, the primer was elongated to nt 4501 due to strong primer dimer and hairpin formation when ending at 4495

3.2 METHODS

3.2.1 QUALITATIVE PCR

Qualitative PCR was used for two purposes in this thesis: i) to clone the *VP2* gene or other regions from the BuV1-, BuV2-, BuV3-, TuV- and CuV-positive patient samples to serve as controls in qPCR as well as to be used in creating the recombinant baculoviruses for VLP expression (Studies I-IV), and ii) to verify the positive qPCR results and to find out the genotype or strain of the virus in the sample (Studies I and IV). These PCRs were carried out with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) with HF or GC buffer according to the manufacturer's specific instructions for Phusion DNA polymerase.

In study I, the BuV qPCR-positive samples were re-amplified either with BuV1 and 2 VP2 primers or with primers BuV1-2 NS fwd KpnI and BuV NS 1kb rev (Table 2).

In study IV, the CuV qPCR-positive tissue samples were re-amplified with CuV 583 fwd and CuV rev or with CuV 230 fwd and CuV rev (Table 2).

3.2.2 QUANTITATIVE PCR

During all qPCR experiments as well as qualitative PCR experiments from patient samples, strict precautions were taken to avoid contamination. Separate rooms were used for preparing the master mixes, handling and extracting DNA from sample material, and preparing the plasmid controls and pipetting the plasmid templates. Aerosol resistant filter tips were used throughout and all runs included H_2O as negative template control.

3.2.2.1 Primer and probe design for qPCR (I,IV)

To design primers and probes for quantitative PCR for BuV, TuV and CuV, all respective sequences available in GenBank were aligned with Clustal W or Clustal Omega, available at the European Bioinformatics Institute webpage (https://www.ebi.ac.uk/services). To obtain sequences for BuV singleplex design, GenBank was accessed 1st Oct 2012, and for TuV and CuV singleplex as well as multiplex, GenBank was accessed 1st Jun 2016.

The conserved and variable regions between the viruses were manually searched for with BioEdit[160] and potential primer and probe sequences were analyzed *in silico* with NetPrimer software (PREMIER Biosoft) and Amplify4 software. The Basic Local Alignment Search Tool (BLAST, The National Center for Biotechnology) was used to see any undesired matching to human or other sequences.

For BuV, the qPCR amplicon was designed to the NS gene region, which is nearly identical between the different BuV types. For TuV and CuV, the qPCR amplicon was designed to the VP2 region of the virus (see also Fig. 17).

3.2.2.2 BuV qPCR (I)

A ten-fold dilution series of BuV1 plasmid containing a 3.5-kb piece of the BuV1 genome, ranging from the NS to VP region, was used in the qPCR set up and as a quantification standard in the final assay. During set up, the sensitivity and specificity of the assay were tested also with the BuV2 plasmid (containing an 830-nt piece of the BuV2 *NS* gene) as well as with i) 500 ng human DNA/reaction (HEK 293 cell line), ii) DNA extracted from human fecal samples spiked with 500 copies of BuV1 plasmid, and iii) high copy number plasmids of other human parvoviruses (B19V, HBoV1 and PARV4).

The final 25-µl reactions consisted of 1x Maxima Probe Master Mix (Thermo Fisher Scientific) with 30 nM ROX reference dye, 0.5 µM of both BuV primers, 0.2 µM of BuV NS1 probe, and 5 µl template. The BuV qPCR was set up with the Stratagene Mx3005P machine (Stratagene) with the following program: initial denaturation of 10 min at 95°C followed by 45 cycles of 15 s at 95°C and 1 min at 60°C. Of note, later on the protocol was tested and used also with the AriaMx Real-Time PCR system (Agilent Technologies). Primers and the probe are listed in Table 2.

3.2.2.3 TuV and CuV singleplex and BuV-TuV-CuV multiplex qPCR (IV)

A ten-fold dilution series of TuV and CuV plasmids containing the *VP2* genes of the respective virus were used to set up qPCRs and as a quantification standard in the final assay. During set up, sensitivity and specificity of the assay were tested as singleplex and multiplex with plasmids, with and without 450ng human DNA/reaction (HEK 293, HeLa, hFSF, HepG2 or A549 cell line), and high copy number plasmids of other human parvo- or polyomaviruses (B19V, HBoV1-4, and PARV4 and a mix of 13 human polyomaviruses).

For TuV and CuV qPCRs, the reaction set up was similar to that of BuV qPCR, except that ROX was excluded from the master mix due the use of this channel for measuring TxRed fluorescence of the TuV probe (Table 2). In addition, doubled concentrations of the degenerate CuV fwd primer and CuV probe were used.

The final multiplex BuV-TuV-CuV-qPCR reaction consisted of 1x Maxima Probe Master Mix (without ROX), 0.5 µM of each primer (1 µM for CuV fwd), 0.2 µM of each probe (0.4 µM for CuV), 5 µl template, and H_2O to a final volume of 25 µl. All primers and probes are listed in Table 2.

The TuV and CuV singleplex and the BuV-TuV-CuV multiplex assays were set up with the AriaMx Real-Time PCR system with the following program: 10 min at 95°C followed by 45 cycles of 15 s at 95°C and 1 min at 62°C. The temperature in the combined annealing and extension step was increased from 60°C to 62°C to ensure the specificity of the assay.

3.2.3 DNA EXTRACTION (I-IV)

All DNA extractions were made according to the manufacturer's instructions unless otherwise indicated.

With cohort 1, the DNA was extracted either from 900 µl of 10% fecal suspension with the NucliSens EasyMag nucleic acid extraction protocol (bioMerieux) using an automated Tecan pipetting robot system (Tecan), or from 200 µl of a 10% fecal suspension with the MagnaPure LC Total nucleic acid isolation kit (Roche Applied Sciences).

With cohort 2, the DNA was extracted from 140 ul 10% fecal suspension and from nasal swabs suspended and mixed in UTM-RT Mini tubes (Copan) using the QIAamp Viral RNA Mini Kit (Qiagen)[127].

With cohorts 9, 10 and 11, DNA from freshly frozen tissue samples (4-5 mm in diameter) was extracted with the QIAamp DNA Mini Kit or AllPrep DNA/RNA/miRNA Universal Kit (Qiagen). DNA from FFPE tissue samples (1–20 × 10 µm sections, depending on the availability and the amount of tissue in the blocks) was extracted with QIAamp DNA FFPE Tissue Kit (Qiagen) with the following modifications: i) tissue sections were deparaffinized twice with xylene (Sigma-Aldrich), ii) the tissues were lysed by incubating overnight in 180 µl ATL buffer and 20 µl proteinase K at 56°C, and iii) DNA was eluted twice, first round with 40 µl and second round with 20 µl ATE elution buffer. DNA from serum (200 µl) was extracted with the QIAamp DNA Blood Mini Kit (Qiagen).

3.2.4 VIRUS-LIKE PARTICLE PRODUCTION (II-IV)

The VP2-pFastBacDual plasmids were transformed into the *E. coli* host strain, DH10Bac™, containing a parent bacmid and helper vector (Bac-to-Bac Baculovirus Expression System, Invitrogen). By using the site-specific transposition properties of the Tn7 transposon in the pFastBacDual vector expression cassette, the entire cassette was transposed site-specifically into the bacmid.

The recombinant bacmid DNA was extracted with alkaline lysis, and the bacmid DNA was transfected into *Spodoptera Frugiperda* (Sf9) cells with Cellfectin transfection reagent (Invitrogen) to generate a recombinant baculovirus. The obtained baculovirus stock was amplified in *Trichoplusia ni* (HighFive) insect cells three times according to the manufacturer's instructions.

For VLP expression, HighFive insect cells were infected with baculovirus (P3 virus stock) and after 5 days of incubation in +27°C, the cells were collected by centrifugation at 4300rpm for 5 min. The cells were re-suspended to 20 mM Tris-HCl, containing Complete-EDTA-free protease inhibitor cocktail (Roche Applied Sciences) and frozen at -20°C.

To free the VLPs from the insect cells, the cell suspension was incubated with 0.5% deoxycholate (Sigma-Aldrich) at +37°C for 30 min with 600 rpm shaking. The treated cells were cooled down on ice and sonicated 8 times with 10 sec sonication/10 sec interval –scheme using 50% amplitude in the sonication phase. After centrifugation of the lysate at 16100 g for 30 min at +4°C, the supernatant was collected, and the sonication procedure was repeated for the remaining pellet. The supernatants from the two sonication rounds were loaded onto a CsCl double cushion with densities of 1.52 g/cm^3 and 1.22 g/cm^3 in TE buffer (10 mM Tris-HCl (pH 7.8), 1 mM EDTA). Ultracentrifugation

was performed at 65 000 g for 4 h or overnight at 10°C. Fractions were collected and those showing pure VP2 bands on SDS-PAGE were dialyzed 3 times against PBS. The purified VLPs were viewed under electron microscopy.

3.2.5 SEROLOGICAL ANALYSIS (II-IV)

All IgG EIAs were done with biotinylated antigens. The purified VLPs as well as the sonicated HighFive insect cell lysate serving as control antigen were biotinylated with EZ-Link Sulfo-NHS-LC-Biotin kit (Thermo Fischer Scientific) according to manufacturer's instructions.

Different plastics and buffers as well as various reagent concentrations were tested during set up.

3.2.5.1 IgG EIA

BuV1-3-TuV-IgG EIA was set up and used as a panel in which each sample was analyzed for each of the virus antigens and the insect cell control antigen in adjacent wells (Study II). The antigens also controlled each other regarding the possible false positive reactivity towards the plastics or other reagents. The CuV antigen was added to this panel in Study III, and the entire panel was used also in Study IV.

Briefly, the final IgG EIA assay was the following: streptavidin-coated plates (Kaivogen) were incubated with 80 ng of biotinylated antigen per well. After post-coating with the Diluent (Labsystems diagnostics) 3 x 10 min, serum samples diluted 1:200 in Red buffer (Kaivogen) were applied to wells and incubated for 1 hour in room temperature with 400 rpm shaking. Wells were washed 4 x with PBST and horseradish-peroxidase-conjugated antihuman IgG (DAKO) was applied and incubated for 1 hour in room temperature with 400 rpm shaking. After the final washes of 4 x PBST, the bound IgG was detected with 3-3',5-5'-tetramethylbenzidine (TMB, Dako in II and III, Sigma-Aldrich in IV). The reaction was stopped with 0.5 M H_2SO_4 and the ODs were measured with Multiscan EX (Thermo Fisher Scientific) at 450 nm wavelength.

To interpret the sample as IgG positive or negative, no classical cut-off was defined, i.e. using the results and standard deviations from known negative samples. Instead, all samples showing of OD≥0.1 were confirmed with competition EIA (see next section), and the final positive IgG result was always based on competition assay result. The OD≥0.1 – value was decided from the information of the early experiments with BuV1 and BuV2 IgG EIA: samples showing OD<0.1 were not reliably repeated in the competition assay and such values have not been considered positive in any other parvoviral IgG EIA in our laboratory before (Lea Hedman, personal communication).

3.2.5.2 IgG EIA competition assay

To study possible cross-reaction and to distinguish those from true reactivity, as well as confirming the IgG-positive reactions for these novel viruses, a competition assay was used as previously described for human bocaviruses[103,105]. For all samples showing OD≥0.1 for any virus antigen, the EIA was repeated with homologous and heterologous blocking. The IgG EIA protocol for competition assay was otherwise identical to the normal IgG EIA, but the serum sample was incubated with the viral VP2-VLPs 20 µg/ml or HighFive insect cell lysate 0.5 mg/ml for 1.5 hours at +4°C prior to adding the sample to the well.

Initially all samples showing reactivity were blocked separately with each of the virus antigens as well as with the insect cell lysate (Study II); later on, only the homologous antigen, the antigen of the phylogenetically closest virus and the antigen of one other protoparvovirus were used in the blocking (Studies III and IV). For example, a sample showing BuV1 reactivity was blocked with BuV1, BuV3 (the closest virus), and either with BuV2 or CuV. In case of unclear results, the blocking was repeated, and other antigens were included to the blocking if needed. All TuV IgG reactivities in Studies III and IV, were tested and blocked with TuV antigens from two different batches as well as at least with two BuV antigens and insect cells. Of note, the routine blocking with insect cells was discontinued in Studies III and IV and used only occasionally since the VLP purity was increased and false positive reactions due to insect cells became rare.

In general, the competition assay result was interpreted as positive, when homologous blocking removed all reactivity, and no or only little heterologous blocking occurred.

For BuV1, BuV3 and TuV, the competition assay result interpretation criteria stayed the same in all studies as described in Study II (the first serological article): "*The overall BuV-IgG result after competition was considered positive when no heterologous blocking occurred and the residual OD fell upon homologous blocking (i) below OD 0.2, when the non-blocked OD was > 0.5; (ii) below OD 0.15, when the non-blocked OD was 0.3–0.5; and (iii) below 50% of the non-blocked OD, when the non-blocked ODs were low, 0.1–0.3*".

After including CuV to the IgG EIA panel, the competition assay result interpretation criteria needed to be updated for BuV2 and CuV due to cross-reactions. A sample was considered CuV or BuV2 IgG positive when homologous antigen blocked the IgG reactivity completely, the heterologous (BuV2 or CuV) only partially, and more distant BuV1 or BuV3 antigen did not block at all.

3.2.6 STATISTICAL ANALYSIS

Statistical analysis was performed either in the OpenEpi software (https://www.OpenEpi.com) with test statistics (mid p-exact value) and 2 x 2 tables (study III), or with R, version 3.4.3 (R Development Core Team) for Mann-Whitney U, and Pearson's χ^2 (study IV). A 2-tailed p value <0.05 was considered statistically significant.

4 RESULTS AND DISCUSSION

4.1 BUV DNA DETECTION IN FINLAND (I, II)

Introduction. In 2012 Phan and co-workers described a new parvovirus, BuV, in a fecal sample of a diarrheic child in Burkina Faso, Africa. In the same study, three more samples from Burkina Faso as well as one from Tunisia were shown by PCR to harbor the DNA of this new virus.

To study the epidemiology and occurrence of BuV in Finland, a sensitive real-time qPCR was designed for detecting and quantifying the viral DNA. In the first study, fecal samples from a cohort of GE patients (n=629) of all ages in Finland were analyzed for BuV DNA, and in the following study, fecal and nasal swab samples from children (n=955) with AGE, ARTI or both symptoms were analyzed to further study the occurrence of this virus.

BuV qPCR design and validation. The analysis of the alignment of four BuV sequences available in GenBank at the time showed conserved areas in the NS region. After manual analysis and *in silico* evaluation, a hydrolysis probe-based qPCR with a 126-nt amplicon was designed. Within the primer areas, all four BuV sequences were identical, and within the probe area, a single nucleotide difference between BuV1 and BuV2 types was covered with a degenerative nucleotide. The commonly used FAM was selected as the fluorescent dye to the probe, and Black Hole Quencher-1 (BHQ1), a non-fluorescent chromophore, was used for quenching the fluorescent signal. The NetPrimer and Amplify4 *in silico* analyses revealed no significant dimers (primer-primer or primer-probe) or hairpin formation.

The analytical sensitivity of the BuV qPCR was 5-10 copies per reaction with both BuV1 and BuV2 plasmids (see 1.4.3 plasmids), and 500 ng of human DNA did not affect the BuV quantification or cause false positives. No false amplification was seen with high copy-number plasmids of the other human parvoviruses, B19V, PARV4 and HBoV1.

At the time when Study I was published, a third genotype of BuV, BuV3, was described[129]. Consequently, a question was raised, whether the developed primers and probe matched the novel BuV3 sequences. In silico analysis demonstrated single nucleotide mismatches between each of the oligos and the BuV3 sequences, however, none of the BuV3 sequences harbored two or all three of the observed mismatches, and therefore, these mismatches were not considered to cause fundamental detection or quantification problems. Of note, later on we analyzed BuV3 DNA-positive fecal samples with the qPCR assay and were able to detect and quantify the BuV3 DNA in the sample even from low copy-number samples (unpublished data).

BuV DNA prevalence among GE patients in Finland (Study I). Seven fecal samples harbored BuV DNA (7/629, 1.1%), of which four were from the bacterial cohort

(4/243, 1.6%) and three from the viral cohort (3/386, 0.8%). All positive samples contained low viral DNA quantities ranging from 1.9E+03 to 3.2E+04 copies per ml of fecal suspension. Despite low copy numbers, the entire 1.7 kb VP2 was amplified from five samples and partial NS (~1 kb) from one, and sequencing showed that all amplicons were most similar to BuV genotype 1 (Technical Appendix Fig. in Study I). One sample was positive only by the short BuV qPCR (repeatedly), but no longer sequence could be obtained. This was most likely due to the very low viral quantity in the sample: 1.9E+03 per ml of fecal suspension, correlating to 9.6E+00 copies per reaction, which was the lowest quantity detected among the seven BuV DNA-positive samples.

In contrast to the other publications at the time, all samples from children were BuV-DNA negative, and the seven samples harboring BuV DNA were from adults (median age 53 y, range 21–89 y). Overall, the analyzed samples were obtained from patients of all ages (median 51.5 y, range 0-99 y), however, the majority of the samples were from adults. Altogether there were only 81 samples from children below the age of 13 years and very few from teenagers, and the smaller sample number may have been one reason why no BuV DNA was detected among children. The age distribution of the patients is depicted in Figure 13 (unpublished data).

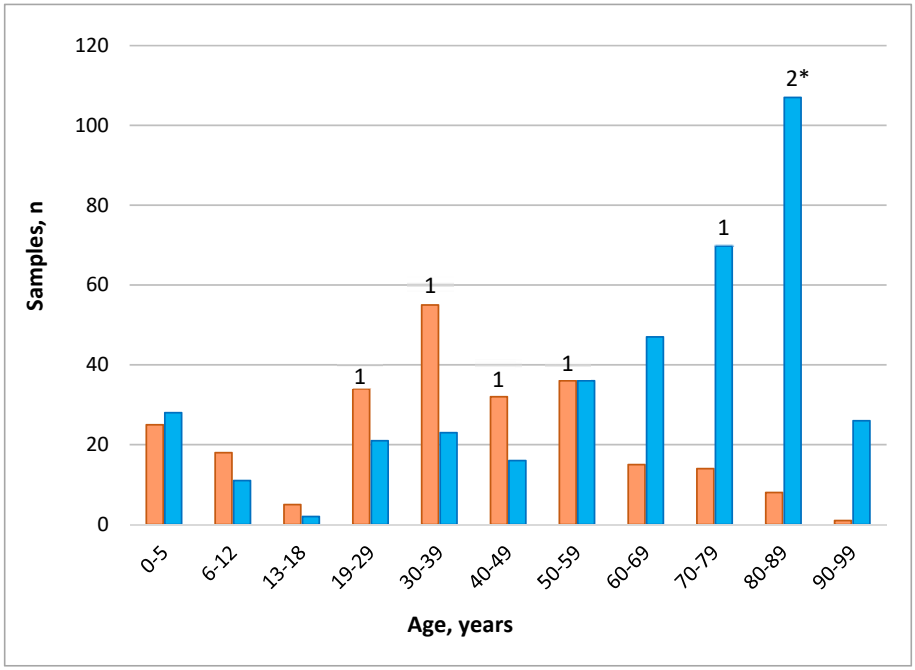


Figure 13. Age distribution of the gastroenteritis patients (n=629) in Finland (Study I). The orange columns represent the bacteriological cohort (n=243) and the blue columns the virological cohort (n=386). The numbers on top of the columns indicate the numbers of BuV DNA-positive samples in each age group and sample cohort. *, both samples were from the same patient

Two of the three positive samples in the viral cohort came from the same patient, an 89-year-old lady living in western Finland, and the samples were taken 4 days apart. The second sample showed a higher BuV copy number than the first, 3.2E+04 vs. 3.6E+03, respectively. The higher viral load in the latter sample as well as the presence of virus in two samples taken four days apart could imply that the virus was not derived from food sources. However, as norovirus was detected in the second sample (the first sample being norovirus negative), it might be that the GE symptoms were caused by norovirus instead of BuV. The other five BuV DNA-positive patients were negative for the GE pathogens tested in routine diagnostics.

BuV DNA prevalence in fecal and nasal swab samples of children in Finland (Study II). To study the BuV occurrence among Finnish children as well as in other sample types, fecal and nasal swab samples from a large pediatric cohort were analyzed. Of the 955 children, three (0.3%) had BuV DNA in the feces and one (0.1%) in the nasal swab sample. The BuV DNA quantity was very low in all four samples ranging from 8.5E+02 to 1.6E+03 copies per ml fecal suspension and 4.6E+03 copies per ml nasal swab medium. The children with BuV DNA in their fecal samples were between 10 months and 4.5 years of age, and the swab-positive child was 20 months of age. All four patients suffered from acute gastroenteritis, and a known GE pathogen was detected in every case. Two children had norovirus, one had rotavirus and HBoV2, and the child with BuV DNA in the nasal swab was positive for norovirus in the feces as well as had a serologically confirmed acute HBoV2 infection. However, the fecal sample of the child remained BuV-DNA negative, and no BuV-IgG antibodies were detected in his serum.

These results showed that BuV can also be found in Finnish children. Although BuV was detected only in children with acute GE symptoms, and all tested 545 non-diarrheic children were BuV-DNA negative, a known GE pathogen was detected in each of the BuV DNA-positive cases and was most likely the etiological agent of the GE symptoms in these children. In addition, the study did not include healthy children without symptoms of an infection, and such individuals with identical sampling times should be analyzed to give a better understanding of the role of BuV in gastroenteritis and other diseases.

Summary of BuV DNA studies. These two studies showed that BuV DNA can be detected in Finland both in adults and in children, although with low prevalence and low quantities, which is in line with the other publications (Table 3). Study I was among the first articles reporting BuV DNA outside Africa as well as the first one reporting BuV in Europe and in adults.

Currently, the presence of BuV DNA in human fecal samples has been reported in 17 studies around the world (including Studies I and II): in nine the analyses were done with PCR (nested, quantitative or real-time) and in eight with NGS. Of note, the results from one NGS study detecting BuV3 reads in Ethiopia has been published in two separate articles[161,162], although here and in Table 3 they are counted as one study.

In the larger, PCR-based studies including >200 study subjects (n=8, Table 3), BuV DNA has been detected in 0.3-1.7% of the fecal samples. A higher BuV-DNA prevalence has been observed only in the original BuV article reporting BuV DNA by PCR in 4.1% (4/98) of the children in Burkina Faso[34], and in two NGS articles reporting BuV reads in 6.7% (4/60) of healthy Amerindian children in isolated Amazonian villages in Venezuela[135] and in 8.3% (1/12) of control children in an IBD study in the USA[137]. Interestingly, even with the growing number of publications, the DNA of BuV2 has still not been reported in any study since the original discovery.

The BuV DNA detected in the nasal swab in Study II is to date the only case describing BuV DNA in another human sample type than feces. The studies have, however, either focused on analyzing fecal samples and environmental water or reporting reads of BuV as one of the NGS findings in a large virome analysis of fecal samples. Overall, there are only three publications besides Study II reporting BuV PCR analyses of other human sample types than feces: a study with cerebrospinal fluid (CSF) samples[163], a study with biopsies from head and neck cancers[164] and Study IV (see also Chapter 4.4). In all three studies no BuV DNA was detected.

In Studies I and II, all samples harboring BuV DNA were collected during winter/spring months. The samples in the children's cohort (Cohort 2) were collected year-round, whereas with Cohort 1, the samples were collected between October and May. It is well known that many other viruses cause epidemics during certain time of the year, and sufficient sampling, including the sampling time span, should be taken into consideration when planning the forthcoming studies. Currently, most BuV studies including GE patients and controls, the samples in the control group have not been collected at the same time period than the actual cases, although the ages of the cases and controls might be well matched.

Regarding the possible association of BuV and gastroenteritis, no strong evidence has been obtained although most of the BuV DNA-positive individuals have had gastroenteritis. Recently BuV DNA has been detected also in non-diarrheal individuals[135–137,165], many BuV DNA-positive individuals have further had a known GE pathogen in the feces and all reported viral loads have been low. Regardless of GE, BuV could cause symptoms of other types, and the detected BuV DNA in feces could originate from previous or current infections unrelated with GE or prolonged virus secretion, which is a common phenomenon with HBoVs[101,128,166].

Table 3. ***Studies describing analysis of BuV DNA.*** *The table includes all studies (Jan 2020) reporting the use on BuV PCR, both with and without BuV DNA findings, as well as all studies reporting NGS reads of BuV. The table is presented in chronological order.*

Study	Sample type	Country	Age (range)	Sampling time	n	Positive (%)	Symptoms	Detection method	Other pathogens in BuV pos samples	BuV genotype, other info
Phan 2012[34]	**Feces**	**Burkina Faso**	<5 y	Nov 2008 - Feb 2010	**98**	**4 (4.1%)**	GE	NGS, nPCR	1 with SaV*, HPeV*; 1 with HBoV*; 1 with NoV*, RVA*, HCosV*; BuV2 sample neg for other viruses*	**Three BuV1, one BuV2**
	Feces	**Tunisia**	"Children"	Unknown	**63**	**1 (1.6%)**	NPAFP	nPCR	Unknown	Type unknown
	Feces	**Chile**	"Children"	Unknown	**100**	**0 (0.0%)**	GE	nPCR	Unknown	-
Yahiro 2014[129]	**Feces**	**Bhutan**	<5 y	Feb 2010 - Jan 2012	**393**	**3 (0.8%)**	GE	nPCR	1 with NoV; 2 neg for enteric viruses	**All BuV3**
Study I	**Feces**	**Finland**	Median 51.5 y (0-99 y)	Oct 2012 - Mar 2013; Apr-May 2013	**629**	**7 (1.1%)**	GE	qPCR	1 with NoV; 2 neg for viruses, bacteria not tested; 4 neg for GE bact., viruses not tested	**Six BuV1, one unknown**
Smits 2014[167]	**Feces**	**the Netherlands**	Median 47 y (0-97 y)	2005-2009	**27**	**1 (3.7%)**	GE	NGS	None	**BuV3**
Phan 2014[35]	**Feces**	**Tunisia**	Mean 18.7 mo (7 days–96 mo)	Unknown	**180**	**unknown**	GE	NGS	Unknown	**168 reads of BuV, type unknown**
Chieo-chansin 2015[168]	**Feces**	**Thailand**	Median unknown (0-97 y)	Jan 2009 - Apr 2014	**1495**	**4 (0.3%)**	GE	nPCR	Child: HPeV 1A; adults neg for enteric viruses	**All BuV1.** BuV pos: 1 child (1 y), 3 adults (19, 48, 90 y)
	Feces	**Thailand**	Median unknown (0-39 y)	Feb 2010 - Jul 2014	**741**	**0**	HFMD	nPCR		-
Altay 2015[130]	**Feces**	**Turkey**	Mean 19.5 mo (1-60 mo)	Sep 2004 - Jun 2011	**583**	**8 (1.4%)**	GE	nPCR	1 with NoV; 1 with NoV, HBoV2; 1 with HBoV3; 5 neg for enteric viruses	**Six BuV3, two unknown**
	Feces	**Turkey**	Mean 17.3 mo (unknown)	Feb-Sep 2013	**148**	**0**	Healthy	nPCR		Age matched controls, but sampling time 1.5-2 y later
Huang 2015[169]	**Feces**	**China, Beijing**	Median 35 y (1 mo-85 y)	2010-2014	**520**	**9 (1.7%)**	GE	Real-time PCR	1 with SaV; 1 with NoV; 7 neg for enteric viruses	**Four BuV1** (all adults), **five BuV3** (3 children, 2 adults)
			"Children"	Unknown	**76**	**0**	Non-GE	Real-time PCR	-	-
	Feces	**China, Chongqing**	Median 10 mo (1 d- 14 y)	2010-2013	**1357**	**0**	GE	Real-time PCR	-	-
			"Children"	Unknown	**345**	**0**	Non-GE	Real-time PCR	-	-

Study	Sample type	Country	Age (range)	Sampling time	n	Positive (%)	Symptoms	Detection method	Other pathogens in BuV pos samples	BuV genotype, other info
Ayouni 2016[170]	**Feces**	**Tunisia**	Median 7.0 mo (0.5-60 mo)	Oct 2010 - Mar 2012	**203**	**2 (1.0%)**	GE	nPCR	1 with NoV, AdV-6; 1 with RVA, AdV-2	**Both BuV1**
Study II	**Feces**	**Finland**	Median 14 mo (6 d-15.6 y)	Sep 2009 - Aug 2011	**172**	**2 (1.2%)**	GE	qPCR	1 with NoV; 1 with RVA, HBoV2	**BuV types unknown**; corresponding swabs BuV neg
					545	**0**	ARTI	qPCR	-	-
					238	**1 (0.4%)**	GE & ARTI	qPCR	NoV	**BuV type unknown**; corresponding swab BuV neg
	Nasal swab	**Finland**	Median 14 mo (6 d-15.6 y)	Sep 2009 - Aug 2011	**172**	**0**	GE	qPCR		-
					545	**0**	ARTI	qPCR		-
					238	**1 (0.4%)**	GE & ARTI	qPCR	Swab neg for HBoV and CoV (feces with NoV, HBoV2)	**BuV type unknown**; corresponding feces BuV neg; Acute HBoV2 diagnosed serologically
Altay 2017[163]	**CSF**	**Turkey**	Median 32 y (0-96 y)	Oct 2011 - Apr 2015	**126**	**0**	Fever/CNS infection#	nPCR	-	-
Altan 2017[171]	**Feces**	**Peru**	Unknown	Unknown	**300**	**unknown**	GE	NGS	Unknown	**NGS reads of BuV3**
Altan 2018[161] Aiemjoy 2019[162]	**Feces**	**Ethiopia**	Median 36 mo (0-5 y)	Apr 2016	**269 samples, 29 pools**	**2 pools**	Unknown	NGS	Unknown	**Two short contigs of BuV3**. One in pool of watery stool and one in smooth stool
Siqueira 2018[135]	**Feces**	**Venezuela, 3 isolated villages**	Mean 3 or 4 y (depending on the village)	2015	**60 (20 per village)**	**4 (6.7%)**	Healthy	NGS	Unknown	**NGS reads of BuV3,** altogether in 4 children from two isolated villages. Also CuV reads in two children, see Table 10.
		Venezuela, Caracas	Mean 4 y	2015	**20**	**0**	Healthy	NGS	-	
Thong-prachum 2018[172]	**Raw sewage**	**Japan**	NA	Jul 2015 - Jun 2016	**12**	**0**	NA	nPCR	-	Other human viruses: RVA, NoV, HBoV2 in 12/12 samples; AstV, SaV, AiV 11/12; AdV-41 8/12; SalV, EV 7/12; HPeV-1 6/12; SAFV 2/12; RVB, RVC, HAV, HEV, RoV, HCosV 0/12
Guerrero-Latorre 2018[173]	**River water**	**Ecuador**	NA	Jun 2017	**3**	**3 (100%)**	NA	NGS	Unknown	**BuV1**

Study	Sample type	Country	Age (range)	Sampling time	n	Positive (%)	Symptoms	Detection method	Other pathogens in BuV pos samples	BuV genotype, other info
Dickinson 2019[164]	**JNA tissue**	**Finland**	Mean 22.7 y (17-33 y)	Unknown	**7**	**0**	JNA	Multiplex qPCR	-	Neg for BuV, TuV, CuV, B19V, HBoV1-4 and 12 HPyVs. One patient had MCPyV.
	Cancer and healthy tissues	**Finland**	Mean 61.2 y (31-86 y)	Unknown	10 cancer, 8 hlty	**0**	OPSCC/ OSCC	Multiplex qPCR	-	Neg for BuV, TuV, CuV, HBoV1-4 and 12 HPyVs. 1 patient had MCPyV, B19V in 7/10 patients.
Fernandes 2019[137]	**Feces**	**USA**	Mean 13.9 y (5-17 y)	Jan - Jun 2015	**7**	**0**	Crohn's disease	NGS	-	Diagnosis at least 6 months before
	Feces	**USA**	Mean 14.6 y (5-17 y)	Jan - Jun 2016	**5**	**0**	Ulcerative colitis	NGS	-	Diagnosis at least 6 months before
	Feces	**USA**	Mean 10.3 y (5-17 y)	Jan - Jun 2017	**12**	**1 (8.3%)**	GE clinic patients§	NGS	Unknown	**15 NGS reads of BuV, type not specified**
Study IV	**Skin biopsies**	**Finland**	Median 64 y (32-89 y)	Unknown	**25**	**0**	CTCL	Multiplex qPCR	-	CuV in 4/25 (16.0%), TuV neg
	Skin biopsies	**Finland**	Median 62 y (22-83 y)	Unknown	**136**	**0**	Transplant	Multiplex qPCR	-	CuV in 4/136 (2.9%), TuV neg
	Skin biopsies	**Finland**	Median 43 y (18-67 y)	Unknown	**98**	**0**	Healthy	Multiplex qPCR	-	Neg for CuV and TuV
Nantachit 2019[174]	**Canal water**	**Thailand**	NA	Nov 2016 - Jul 2018	**125**	**1 (0.8%)**	NA	nPCR	Unknown	**BuV1**
Tokarz 2019[136]	**Feces**	**North America**	Mean 13 y (4-17 y)	Jul 2012 - Apr 2015	**70**	**0**	Ulcerative colitis	NGS	-	NGS with VirCapSeq-VERT enrichment
	Feces	**North America**	Unknown	Jul 2012 - Apr 2015	**70**	**1 (1.4%)**	IBD	NGS	Unknown	**NGS reads of BuV, type not specified.** NGS with VirCapSeq-VERT enrichment
Okitsu 2019[165]	**Feces**	**Bangladesh**	73 +/- 4 mo (66-81 mo)	Jun - Oct 2016	**227**	**1 (0.4%)**	Non-GE¤	nPCR	HCosV	**BuV type unknown.**
Mohammad 2020[175]	**Feces**	**Kuwait**	Median 2 y (1 mo - 10 y)	Jan - Dec 2017	**84**	**1 (1.2%)**	GE	NGS	AdV, HCosV	**One NGS read of BuV3**

*, NGS reads; #, patients clinically diagnosed with febrile disease and/or CNS infections of presumed viral etiology; §, healthy, non-IBD subjects, 11/12 had either constipation or abdominal pain; ¤, non-diarrheal samples from children enrolled to a study of association between asthma and parasitic infections

AdV, adenovirus; AiV, aichivirus; ARTI, acute respiratory tract infection; AstV, astrovirus; CNS, central nervous system; CSF, cerebrospinal fluid; CoV, coronavirus; CTCL, cutaneous T-cell lymphoma; d, day; EV, enterovirus; GE, gastroenteritis; HAV, hepatitis A virus; HBoV, human bocavirus; HCosV, human cosavirus; HEV, hepatitis E virus; HFMD, hand, foot, and mouth disease; hlty, healthy; HPeV, human parechovirus; HPyVs, human polyomaviruses; IBD, inflammatory bowel disease; JNA, Juvenile nasopharyngeal angiofibroma; MCPyV, merkel cell polyomavirus; mo, month; NA, not applicable; neg, negative; NPAFP, nonpolio acute flaccid paralysis; NoV, norovirus; nPCR, nested PCR; OPSCC/OSCC, oropharyngeal and oral cavity squamous cell carcinoma; pos, positive; qPCR, quantitative PCR; RoV, rosavirus; RVA, rotavirus A; RVB, rotavirus B; RVC, rotavirus C; SalV, salivirus; SaV, sapo virus; SAFV, saffoldvirus; y, year

4.2 FIRST SEROLOGICAL STUDY OF BUV AND TUV (II)

Introduction. The mere presence of viral DNA in a secretion sample does not necessarily mean that the virus has infected the individual. Viral DNA present in a fecal sample can be derived from food sources and in nasal swab from inhaled air. Especially when the viral quantity in a sample is low, it is difficult to evaluate whether the detected DNA is coming from an infection or from an outside source. Viral antibodies provide an excellent marker of a true infection as well as a good record of past infections.

It has been shown for B19V, HBoV1-4 and PARV4, that the major capsid protein VP2 (or VP3) alone can form the capsid, and such VLPs can be used as an antigen in serological assays[14,176–179]. Therefore, VP2 proteins of BuV1, BuV2, BuV3 and TuV were expressed in insect cell cultures, purified with ultracentrifugation, and used as antigens in an IgG-EIA assay. Using the novel method, BuV1-3 and TuV IgG antibodies were analyzed among adults (n=180) and children (n=228) in Finland to evaluate the IgG seroprevalence of these viruses.

BuV1-3 and TuV VP2 amplification and cloning. The VP2 genes of all four viruses were successfully amplified from the patient samples and cloned to pFastBacDual donor vector with BamHI and SalI restriction sites.

The BuV1-*VP2* gene of 1710 bps was amplified from sample BF.7 (GenBank #JX027295). The obtained sequence had two nucleotide mismatches compared to JX027295, of which one was synonymous and the other caused an amino-acid change in position 421 (threonine to isoleucine). The amplification and cloning were repeated, nevertheless, the nucleotide change remained. As the same nucleotide was obtained in two different amplification rounds indicates that the nucleotide might be truly present in the virus, and the clone was used to create the recombinant baculovirus and BuV1-VP2 VLPs. The VP2 sequence was deposited to GenBank (#KX856937).

The BuV2-*VP2* gene of 1710 bps was amplified from sample BF.39 (GenBank #JX027297). The obtained sequence had a single nucleotide mismatch compared to JX027297, which was synonymous, and thus the amplification was not repeated. The clone was used to create the recombinant baculovirus and BuV2-VP2 VLPs, and the VP2 sequence was deposited to GenBank (#KX856938).

The BuV3-*VP2* gene of 1719 bps was amplified from sample BTN-310 (GenBank #AB847989), and the obtained sequence was identical to AB847989. The clone was used to create the recombinant baculovirus and BuV3-VP2 VLPs, and the VP2 sequence was deposited to GenBank (#KX856939).

The TuV-*VP2* gene of 1698 bps was amplified from sample Tu491 (GenBank #KJ495710), and the obtained sequence was identical to KJ495710. The clone was used to create the recombinant baculovirus and TuV-VP2 VLPs, and the VP2 sequence was deposited to GenBank (#KX856940).

Recombinant BuV1-3 and TuV VP2 protein expression and VLP purification. HighFive insect cells were infected with the recombinant baculoviruses and the SDS-PAGE analysis of the infected cells showed a band located between the 50 and 75-kDa bands of the protein size marker. After purifying the putative VLPs with the CsCl ultracentrifugation, the same band of ~64 kDa for BuVs and ~62 kDa for TuV was present in SDS-PAGE (Fig. 1 in Study II). Electron microscopy of the purified VLPs showed parvoviral-like capsids of ~25 nm in diameter (Fig. 14)

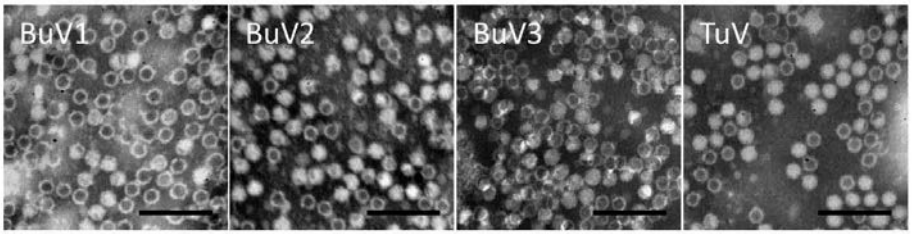


Figure 14. Electron microscopy pictures of purified BuV1, BuV2, BuV3 and TuV-VP2 VLPs. Scale bar 100 nm.

BuV1-3 and TuV IgG EIA assay. An antigen concentration of 80 ng per well was shown to be optimal for all four viruses, and all IgG EIAs could be carried out in the same assay conditions facilitating the use of the four EIAs as a panel. In the IgG panel, a single well per antigen was used for sample screening, and the antigens were controlling each other for example regarding possible unspecific background caused by plastics or other reagents. HighFive insect cell antigen (biotinylated cell lysate) was included to the panel to monitor any possible cross-reactions from insect cell traces left in the purified VLPs.

BuV-IgG prevalence in adults. BuV IgG was detected in ten subjects (10/180, 5.6%) of which four had BuV1 IgG, five BuV2 IgG and one had IgG towards all three BuV types. Interestingly, although 90.7% of the subjects (163/180) were of Finnish descent, half of the BuV IgG-positive subjects originated from India, the Middle East or China resulting in a remarkable difference in BuV IgG seropositivity between Asian (5/12, 41.7%) and native Finnish subjects (5/163, 3.1%). While the result was intriguing and could direct further serological studies towards Asian and Middle Eastern countries, it did not reach statistical significance due to the low number of samples.

Persons with Asian background had moved to Finland as adults and were 26 to 34 years of age during sampling time. The other BuV IgG-positive persons were 22-26-year-olds, except one that was 56 years old. Among the Finnish BuV IgG-positive subjects, all showed low OD values, whereas the subjects with high ODs were from Asia. These included three BuV1 IgG-positive persons from India and the triple positive person was from the Middle East.

To study if the BuV IgG seropositivity is long-lived, similarly to other human parvoviruses, follow-up samples were analyzed from five seropositive persons. The

analysis showed that IgG could be detected for at least 6 years, and in all cases the antibody titre reflected by the OD values remained stable throughout the follow-up period (Table 2 in Study II and Table 4). Overall, these results show that BuV infections elicit strong and long-lasting immune reactions similarly to those of B19V and HBoV1.

BuV-IgG prevalence in children. Of the 955 children with stool and nasal swab samples that were analyzed by BuV qPCR, serum samples were available from 228 children. Seven (3.1%) harbored BuV IgG: one for BuV1, three for BuV2, two for BuV3 and one had both BuV1 and BuV2 IgG. The children with BuV1 and/or BuV2 IgG were 1- to 5-year-olds, whereas the two children with BuV3 IgG were only 7 and 8 months of age, and thus the low-level BuV3 IgG could be of maternal origin. In general, the maternal antibodies are considered to disappear from the child by the age of six months, but without knowing the serological status of the mother, the possibility should be kept in mind.

From the three children with BuV DNA detected in the stool, no serum sample was available, and for the child with BuV DNA in the nasal swab, the corresponding serum sample was BuV-IgG negative. The lack of IgG in the latter child could indicate that either the BuV DNA detected in the nasal swab was not from an infection or that the infection was at a too early phase for IgG to have formed yet. Although this child was suffering from many underlying health issues such as heart problems, asthma and operated cleft palate, a serologically confirmed acute HBoV2 infection[127] indicated that the child had a normally functioning immune system, and a normal immune response towards (parvo)viral infections.

The overall BuV seroprevalence among the Finnish children was the same as among the native Finnish adults. Why the prevalence did not rise with age, is not known. Waning of antibodies can have an effect, however, it was shown that the IgG stayed for at least 6 years without any change in the OD value (OD value reflects the antibody titre in a sample) (Table 2 in Study II). As the circulation of these viruses in Finland seems to be rare, a natural boosting of the antibody titre by repeated exposures to the virus does not occur, thus the antibody level could eventually wane away. All samples with high OD were from individuals who had moved to Finland as adults and most likely had been infected before. It could be hypothesized that these individuals had been living in an environment where these viruses circulate, and they acquired natural stimulation to the immunity and maintenance of the specific antibodies.

BuV competition assay and cross-reactivity. Previously it has been shown for HBoV1-4 that these closely related parvoviruses cross-react in serological assays, and the results need to be verified by a competition assay[105]. The HBoVs differ 11-23% at the VP2 protein level[105], and although the VP2 proteins of the three BuVs differ more (27-35%)[129,133], cross-reactions may occur. Thus, the competition assay was used to distinguish between specific and false reactivity in the IgG EIA as well as study the potential cross-reactivity of the viruses.

The competition-assay results suggested that BuV genotypes represent also serotypes. When BuV IgG-positive samples were blocked with different antigens in separate wells, only the homologous antigen was able to remove the reactivity while the heterologous antigens did not (Table 4 and Table 3 in Study II). The same result was observed with all three BuV types, with samples harboring IgG towards one, two or all three BuV types, and with samples with high or low OD values: each IgG reactivity was blocked only with the homologous antigen (Table 3 in Study II). In addition, the blocking characteristics were identical in the follow-up samples taken up to 6 years apart (Table 4). The lack of cross-reactivity and the presence of specific antibodies towards more than one BuV in a single individual argues against the presence of original antigenic sin between the BuV types. However, properly diagnosed consecutive infections of different BuV types are needed to fully address this question.

Table 4. ***Examples of IgG competition-assay results from adults with follow-up samples****. The samples were taken 6 months (case 2) and 6 years (cases 1 and 5) apart. In each sample only the homologous antigen blocked the reactivity, i.e. BuV1 blocked BuV1 etc. The numbers represent OD values from the IgG EIA and specific blocking is illustrated with bolded numbers. Further examples of competition-assay results are presented in Table 3 of Study II.*

Subject		Case 1						Case 2		Case 5	
Sampling date		4.3.2010			18.2.2016			6.1. 2011	19.7. 2011	19.10. 2009	17.2. 2016
Antigen in the well		**BuV1**	**BuV2**	**BuV3**	**BuV1**	**BuV2**	**BuV3**	**BuV1**	**BuV1**	**BuV1**	**BuV1**
Blocking antigen	**None**	2.588	2.761	2.896	2.683	2.419	2.675	3.095	3.855	0.223	0.249
	BuV1	**0.030**	2.587	2.661	**0.033**	2.356	2.567	**0.035**	**0.029**	**0.058**	**0.095**
	BuV2	2.579	**0.038**	2.722	2.487	**0.032**	2.614	3.063	3.285	0.217	0.256
	BuV3	2.400	2.614	**0.080**	2.532	2.391	**0.088**	3.532	3.466	0.204	0.270
	TuV	2.513	2.720	2.819	2.530	2.424	2.686	3.369	ND	0.196	0.252
	H5	2.638	2.845	2.954	2.653	2.630	2.645	3.702	3.182	0.221	0.269

H5, HighFive insect cells; ND, not done

TuV-IgG prevalence in adults and children, and competition assay results. All adults were TuV-IgG negative, while TuV IgG was detected in one 9-year-old child (1/228, 0.4%). Although the OD value was low (OD 0.31), five times more blocking antigen was needed to remove the reactivity. However, as the other antigens did not clearly block the reactivity, the sample was considered TuV IgG positive (Table 3 in Study II). Nevertheless, it needs to be remembered that there are no positive controls for TuV IgG assay development and therefore the functionality of the assay is not fully known. The IgG EIA assays for other human parvoviruses using the same assay principles than the TuV IgG EIA have been successful before, and therefore the TuV IgG EIA could be hypothesized to function as well.

Insect cell derived cross-reactivity. Despite that the viral antigens showed no cross-reactivity, false positive reactions were occasionally demonstrated, mainly deriving from insect cell traces. From 5 to 6% of the tested samples both in the adult's (9/180) and children's (13/228) cohorts exhibited reactivity to HighFive insect cell control antigen. Therefore, insect cell lysate was used in the competition assay as one of the blocking antigens to further confirm the IgG results.

Of the viral antigens, the BuV2-VP2 VLPs were particularly prone to contain insect cell impurities. Whether a lower expression of BuV2 VP2 in general, or other properties of the capsid, was the main reason for the insect cell background remains to be elucidated. Different variations of the VLP purification method were tested to increase the purity of BuV2-VP2 VLPs, and ultracentrifugation of the VLPs twice was shown to be the best option (data not shown).

In the competition assay the IgG reactivity was considered to be caused by insect cell traces when blocking with the insect cell lysate removed the reactivity and the homologous virus antigen did not or did less so. Of note, also the virus antigen used for blocking contained insect traces as did the biotinylated antigen in the well, and therefore was able to block some of the reactivity caused by the insect cells. An example of the insect cell background and how the blocking assay worked in such cases is presented in Table 5.

Table 5. ***Example of insect cell derived false reactivity to BuV2.*** *The competition-assay results from a sample from a healthy Finnish adult showing false reactivity to BuV2. The sample was interpreted as BuV2 IgG negative as the reactivity was not removed by BuV2, but with insect cells instead. The numbers represent OD values from the IgG EIA, and specific blocking is illustrated with a bolded number.*

	Antigen in the well	BuV2
Blocking antigen	none	0.159
	BuV1	0.168
	BuV2	0.168
	H5 insect cells	**0.052**

H5, HighFive insect cells

4.3 GLOBAL SEROPREVALENCE OF BUV, TUV AND CUV (III)

Introduction. The results from Study II showed that IgG towards all three BuVs can be detected among children and adults, and the three BuVs seemed to be serotypes. The high BuV IgG prevalence detected among the Asian adults compared to the Finnish individuals, suggested that there might be countries and areas where BuVs are more prevalent. This inspired us to study populations from different countries and continents to find endemic areas for the three BuVs.

Besides the three BuVs and TuV, another human protoparvovirus, cutavirus (CuV), was discovered in 2016. Based on the phylogenetic analysis of NS1, CuV was classified as a close relative of BuV, although a different species. However, in the VP2, BuV2 was closer to CuV than it was to BuV1 and BuV3: the VP2 proteins of BuV2 and CuV showed 82% similarity, whereas the VP2s of BuV1 and BuV3 were only 63-65% similar to that of CuV. Although the BuVs did not cross-react in our previous study, such reactions might occur between BuV2 and CuV.

We expressed and purified the CuV-VP2 VLPs and included CuV to the IgG-EIA panel and competition assay. With the IgG EIA panel, plasma or serum samples from altogether 733 adults from Finland, the United States, Iran, Iraq and Kenya, as well as from 107 children from Kenya, were analyzed to determine the BuV, TuV and CuV IgG prevalence in five countries located on four continents.

CuV VP2-VLPs and IgG EIA. The CuV *VP2* gene of 1710 bps was amplified from sample Br337 (GenBank #KT868811). The obtained sequence had two nucleotide mismatches compared to KT868811, however, they were synonymous. The *VP2* gene was cloned to the pFastBacDual donor vector with BamHI and SalI restriction sites, and the clone was used in creating the recombinant baculovirus and CuV-VP2 VLPs as described earlier for BuVs and TuV. The CuV VP2 sequence was deposited to GenBank (#MH127919).

SDS-PAGE analysis of the purified CuV-VP2-VLP yielded a band of ~64 kDa (Fig. 1 in Study III), and electron microscopy showed parvoviral-like capsids of ~25 nm in diameter (Fig. 15).

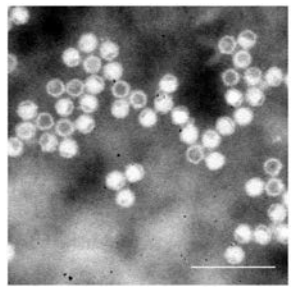


Figure 15. Electron microscopy picture of purified CuV-VP2 VLPs. The scale bar 100 nm.

CuV-VP2 VLPs were biotinylated and used as antigens in the EIA. The same IgG EIA protocol as used for BuVs and TuV was shown to work with CuV as well, and CuV was thus included in the existing human protoparvovirus IgG EIA panel.

Global BuV-IgG seroprevalence in adults. The BuV-IgG prevalence varied greatly between the countries: in Finland and in the USA, BuV IgG was detected in 2-4% of the adults, whereas in Iran, Iraq and Kenya, 56-85% of the adults were seropositive (grey columns in Fig. 16 and Table 2 in Study III). These results indicated that the Middle East and Africa are endemic areas for BuV infections, and in contrast, in Finland and in the USA BuV is not common. The result also confirmed our preliminary finding in Study II that there are large variations in BuV-IgG prevalence between different geographical regions.

The predominant BuV type was BuV1 in the Middle East, and BuV3 in Kenya and in the USA (Fig. 16). Within the endemic areas, the IgG prevalence of the three genotypes were more evenly distributed in Kenya than in the Middle East. In Kenya the predominant BuV3 IgG was present in 65% of all BuV-IgG positives (56 BuV3 IgG positives among 86 BuV seropositives of any type), whereas in the Middle-East the predominant BuV1 IgG was present in 92-95% of the BuV IgG positives (80 BuV1 IgG positives among the 86 BuV seropositives in Iraq and 55 BuV1 IgG positives among the 60 BuV seropositives in Iran) (Fig. 16).

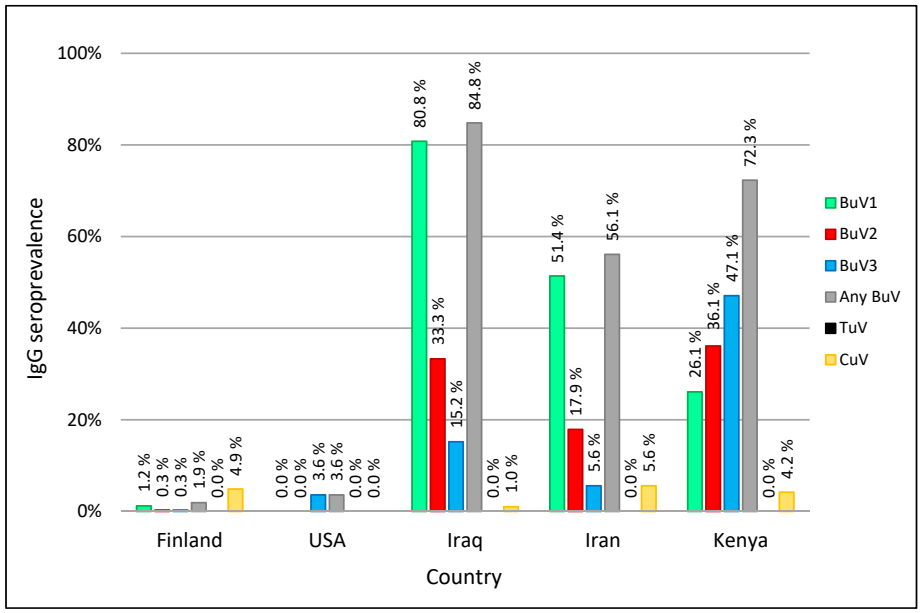


Figure 16. **IgG seroprevalence of BuV1, BuV2, BuV3, TuV and CuV in adults in five countries.** The "any BuV" bar indicates the persons harboring BuV IgG towards one or more BuV types.

BuV2 was the second most prevalent in IgG EIA in all endemic countries. This is an interesting finding as BuV2 DNA has been detected only in a single child to date[34]. All other BuV DNA findings have been of BuV1 or 3 worldwide when genotyping has been possible. As BuV DNA has been searched for mostly in fecal samples, it could be that BuV2 is more likely present in a different sample type for example from the respiratory system or from blood if the patient was viremic.

Individuals that harbored BuV antibodies towards more than one BuV type were encountered most among the Iraqi and Kenyan adults: 30.3% (30/99) of Iraqis and 25.5% (30/119) of Kenyans had antibodies against two BuV genotypes and 7% (7/99) and 5% (6/119) against all three, respectively. In Iran, the double and triple IgG positive persons were less common, 13.1% (14/107) and 1.9% (2/107), respectively. In Finland and in the USA each BuV IgG positive individual had IgG of only one BuV type. Of note, in Finland one veterinarian had both BuV2 and CuV IgG.

In contrast to Study II, some cross-reactions between the three BuVs were observed in the samples from countries with high seroprevalence. The cross-reactions were mostly seen in samples with high OD, and they were easily distinguished from true positivity with the competition assay. Nevertheless, the low-level false reactions demonstrated that all IgG reactivities should be verified with the competition assay for correct result interpretation.

BuV-IgG prevalence among veterinarians in Finland. In Study III, all studied adults from Finland were veterinarians (82%), veterinary students or veterinary nurses. The BuV IgG prevalence was 1.9% (6/324), which was very similar to the prevalence among Finnish adults in Study II (3.1%, 5/163). This suggests that BuV infections are not increased by animal contacts, at least among Finnish veterinarians. In addition, no specific animal contacts were observed among the BuV IgG-positive veterinarians compared to the rest of the cohort.

Age related BuV-IgG prevalence. The BuV seroprevalence among children in Kenya rose according to age: 6/50 (12.0%) of the <5-year-old children had antibodies towards one or more BuV types compared to that of 16/57 (28.1%) in 5-17-year-old children, and 86/119 (72.3%) in adults (>18 y) (<5 y vs 5-17 y, P=0.043, <5 y vs >18 y, P<0.0001, 5-17 y vs >18 y, P<0.0001) (Table 6). The BuV genotype distribution pattern was similar in all

Table 6. ***BuV1, BuV2, BuV3, TuV and CuV IgG seroprevalence according to age in Kenya.*** *The "Any BuV" column presents individuals that harbor IgG towards one or multiple BuV types. The numbers represent the number of positives (%).*

	n	Any BuV	BuV1	BuV2	BuV3	TuV	CuV
<5 years	50	6 (12.0%)	0 (0.0%)	0 (0.0%)	6 (12.0%)	0 (0.0%)	2 (4.0%)
5-17 years	57	16 (28.1%)	3 (5.3%)	4 (7.0%)	14 (24.6%)	0 (0.0%)	0 (0.0%)
>18 years	119	86 (72.3%)	31 (26.1%)	43 (36.1%)	56 (47.1%)	0 (0.0%)	5 (4.2%)

age groups. These results suggest that the age of acquisition of BuV differs from HBoV1. but resembles more that of B19V.

The age-dependent increase of BuV seroprevalence prompted the analysis of the adult cohorts according to age. The cohorts were divided to under and over 40-year-olds, and in Iran a clear difference in BuV seroprevalence was observed: 39% (21/54) of the 18-40-year-olds harbored BuV IgG compared to that of 74% (39/53) in 40-77-year-olds (P=0.0003) (Table 3 in Study III). Such a difference was not seen in Iraq (89% 18-40 y vs 82% 40-60 y) or in Kenya (71% 18-40 y vs 72% 40-88 y). However, when the adult cohorts were divided into under and over 30-year-olds, in Kenya an age-dependent trend was noted (56% [14/35] vs 76% [70/92], P=0.0595), whereas in Iraq the trend was still absent. The differences seen between the under and over 40-year-olds in Iran could be caused by cultural or socioeconomic changes over time or by decreased BuV circulation in the area. In Finland and in the USA the low overall BuV seroprevalence caused any difference to be statistically insignificant.

Global TuV-IgG seroprevalence. TuV IgG was absent in all cohorts. As the TuV is phylogenetically closest to animal protoparvoviruses, it has been speculated whether TuV could actually be an animal virus, for example a rodent virus, instead of a human virus. The viral DNA originally detected in the child's feces could have been derived from food or the environment. In general, parvoviruses are species-specific, i.e. they are causing infections only in a specific host or its close relatives and crossing the species line to another host is very rare. Here, 324 veterinarians from Finland were TuV IgG negative, and thus if TuV infects pets, livestock or fur animals in Finland, the virus did not jump into another host and infect the veterinarians treating the sick animals.

Global CuV-IgG seroprevalence. In contrast to the significant differences in global BuV seroprevalence, the CuV IgG prevalence was generally low in all studied populations (Fig. 16). The highest prevalence was among Iranians (5.6%, 6/107), Finnish veterinarians (4.9%, 16/324) and Kenyan adults (4.2%, 5/119). In the cohort of healthy blood donors from the USA CuV IgG was absent, and in Iraq only one person (1%, 1/99) harbored CuV IgG. Among the Kenyan children, two 3-year-old children had CuV IgG with high OD, while all the other children were CuV-IgG negative (age range 0.5-17 y).

Interestingly, in Iran all six individuals with CuV IgG were over 50 years of age, while in Kenya all five CuV IgG positives were under 40 years of age. In the veterinarian cohort in Finland, the CuV IgG-positive persons were between 29 and 66 years of age, and in Iraq the single CuV IgG-positive individual was 25 years of age. The different age distributions could be explained by rarity of CuV infections in general or by possible differences in the epidemiology of CuV. It could also be that CuV was introduced to Kenya only recently. In addition, the presence of serologically divergent CuV, that cannot be accurately detected by our IgG serology, is possible. More longitudinal serological as well as DNA studies are needed to confirm the variation and possible reason behind the CuV epidemiology.

CuV and BuV2 IgG cross-reactivity. CuV and BuV2 were clearly cross-reactive in the IgG EIA, and in nearly all samples showing reactivity towards one of the antigens, reactivity towards the other was observed as well. Only in samples with low OD the cross-reaction was occasionally absent. Nevertheless, the competition assay was successfully used to distinguish true from false positivity: the true reactivity was blocked completely only with the homologous antigen but was only lowered by the other, whereas the false reactivity was completely removed by both (Table 7). The cross-reactions could be separated reliably in most of the cases, but the results from six unclear cases were excluded from the prevalence calculations. In addition, we encountered three individuals with IgG towards both CuV and BuV2, which suggested that the antibodies towards one virus does not prevent the formation of antibodies towards the other virus. Thus, it seems that the infection of BuV2 or CuV does not cross-protect from infection of the other, however, further studies are needed for more conclusive results.

Table 7. ***Competition assay results from three Kenyan patients showing BuV2 and CuV IgG reactivities.*** *The sample was considered positive, when only the homologous antigen blocked the reactivity completely. For example, with patient 126, the CuV reactivity is blocked completely only with CuV whereas the BuV2 reactivity is blocked with both BuV2 and CuV. The numbers represent OD values from the IgG EIA, and blocking is illustrated with bolded numbers.*

Patient no, age		462, 36 y		242, 20 y		126, 22 y	
Result interpretation		BuV2 and CuV IgG pos		BuV2 IgG pos		CuV IgG pos	
Antigen in the well		**BuV2**	**CuV**	**BuV2**	**CuV**	**BuV2**	**CuV**
Blocking antigen	**none**	2.566	1.898	1.194	0.475	1.715	2.501
	BuV1	2.390	1.797	1.172	0.472	1.598	2.312
	BuV2	**0.081**	0.770	**0.035**	**0.046**	**0.036**	1.484
	CuV	1.582	**0.107**	0.510	**0.044**	**0.031**	**0.028**

y, year

BuV, TuV, and CuV IgG prevalence among HIV-positive individuals. The Kenyan cohorts included HIV-positive persons, 38/119 in the adult cohort and 9/107 in the child cohort. The BuV seroprevalence was very similar between the HIV-positive adults (median age 46 y, range 27–85 y) and HIV-negative adults (median age 38 y, range 18–88 y): 79% (30/38) versus 69% (56/81) had BuV IgG, respectively (P=0.275). When the age ranges of the HIV-positive and -negative adults were matched, the seroprevalences were even more alike; 79% vs 75% (P=0.669), respectively (Table 8). Similarly, the CuV IgG prevalence did not vary between the HIV-positive and -negative adults: 2.6% (1/38) versus 3.4% (4/81) (no age matching, P=0.627), or 2.6% versus 3.3% (2/60), (age matched cohorts, P=0.892). The HIV-positive and -negative children had also similar BuV and CuV IgG seroprevalences, however, the low number of HIV-positive children restricted the statistical comparisons. All Kenyans were TuV IgG-negative.

These results suggested that the infection route of BuV and CuV differs from that of HIV in Kenya, and that HIV infection does not predispose the individual for BuV or CuV infection. With other human parvoviruses, PARV4 has been shown to be transmitted through blood-borne routes similar to HIV and HCV in Europe with the highest risk in intravenous drug users: the PARV4 IgG seroprevalence was 20-95% among risk groups compared to 0-4.7% among healthy non-risk individuals[178–183]. However, in Africa PARV4 is spread differently. For example, in rural areas of Ghana with low HIV prevalence in the population, PARV4 DNA has been detected in blood, nasal swabs and fecal samples of young children indicating fecal oral or respiratory routes of infection[184–186]. In addition, in Burkina Faso, Cameroon, Congo, Botswana and South Africa, the seroprevalence between healthy adults (4-38%) and risk groups (HIV or HCV positives, 19-64%) differs less from each other than in Europe[183,187–189]. Here we did not analyze samples from HIV/HCV-positive patients or other risk groups in Finland or in the USA, and further studies are needed to elucidate, does BuV or CuV have such parenteral infection routes in the Northern hemisphere as PARV4.

In addition, although the antibody results cannot be compared to detection of viral DNA, it could be pointed out that in the study of rhesus macaques experimentally infected with SIV, the DNA of BuV-like virus (WUHARV parvovirus) was detected more often in the feces of SIV-infected animals than in healthy controls (29% vs. 2%, Table 1)[146]. However, these monkeys did not receive any medication for SIV, and developed symptoms and AIDS rapidly shown by the fact that nearly half of the animals needed to be euthanized before week 64 post infection due severe AIDS. In our IgG study the samples from the HIV positive patients were collected at the local HIV clinics, and 43/47 were reported to receive HAART thus indicating that the HIV infection was under control. In all, BuV (or CuV) could cause more infections in AIDS patients, but based on our data on antibody prevalence, the HIV patients do not seem to be a specific risk group.

Table 8. ***BuV, TuV and CuV IgG seroprevalences of HIV-positive and -negative individuals in Kenya****. The age ranges of the HIV-positive and -negative adult cohorts were matched to exclude potential age-related bias.*

Cohort	n	Median age, range	Any BuV	BuV1	BuV2	BuV3	TuV	CuV
Adults, HIV pos	38	46 y, 27-85 y	30 (78.9%*)	10 (26.3%)	18 (47.4%)	19 (50.0%)	0	1 (2.6%#)
Adults, HIV neg	60	45 y, 27-82 y	45 (75.0%*)	17 (28.3%)	21 (35.0%)	30 (50.0%)	0	2 (3.3%#)
Children, HIV pos	9	9 y, 4-17 y	2 (22.2%)	0	0	2 (22.2%)	0	0
Children, HIV neg	98	5 y, 0.5-17 y	20 (20.4%)	3 (3.1%)	4 (4.1%)	18 (18.4%)	0	2 (2.0%)

*, P=0.669; #, P=0.8922; y, year

4.4 CUV IN CTCL AND TRANSPLANT PATIENTS, AND HEALTHY ADULTS (IV)

Introduction. The presence of CuV DNA skin biopsies of CTCL patients was reported already in the original CuV discovery paper, and later on, a Danish group described CuV sequences in skin biopsy from a melanoma patient. CTCL represents a group of T-cell malignancies that arise on the skin, and there are various subtypes of which mycosis fungoides is the most common[190]. The etiology of CTCL is unknown, although several studies have explored and suggested infectious causes (reviewed in [191]).

To study the frequency and possible role of CuV in CTCL and other skin diseases, skin biopsies both from healthy and/or malignant skin from 25 CTCL patients, 136 transplant recipients and 98 healthy adults were analyzed. For CuV DNA detection, a CuV qPCR was developed, and together with a new TuV qPCR, the PCRs were combined with BuV qPCR to create a multiplex qPCR detecting the three human protoparvoviruses in a single run. In addition to the DNA, BuV, TuV and CuV IgG was analyzed from serum samples of 42 CTCL, 124 transplant patients and 78 healthy adults.

CuV and TuV singleplex qPCR and BuV-TuV-CuV multiplex PCR design. For CuV qPCR, an alignment including all known CuV and BuV sequences was manually analyzed to pinpoint regions with similarity among CuVs and variability towards BuVs. After careful in silico and in vitro testing, a primer pair and a probe amplifying a 91-nt piece close to the end of VP2 of CuV were selected (Fig. 17). For TuV, only one sequence was available, and the qPCR was designed based on the VP2 region of the virus. The final TuV qPCR amplified 118 nucleotide long piece of VP2 area (Fig. 17).

For the multiplex qPCR, no dimers between the oligos from the three individual qPCR were observed, neither in silico nor in vitro analysis.

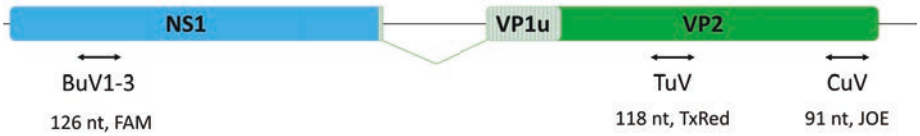


Figure 17. Schematic illustration of BuV-TuV-CuV multiplex qPCR amplicon locations. Amplicon length and fluorophores used in the probes are indicated. Genome based on BuV1, strain BF.96 (GenBank no. JQ918261).

Performance of the qPCRs. The singleplex CuV and TuV qPCRs as well as the BuV-TuV-CuV multiplex PCR amplified the viral sequences in all 20 replicates of 10 copies/µl of plasmid, and the sensitivity remained when 450 ng of human DNA per reaction (90 ng/µl of template) was added to the amplification reaction. For BuV singleplex qPCR, which was set up already in Study I, the sensitivity and specificity testing was repeated here for comparison as well as due to the increased annealing temperature (from 60°C to 62°C), and the results were identical to CuV and TuV qPCRs. Overall, the results indicate

that the analytical sensitivity, or the limit of detection (LOD), of the singleplex and multiplex qPCRs was less or equal to 10 copies/µl, both with and without human DNA.

The LOD was further studied by probit analysis (see Study IV supplementary material for further information). The analysis showed the following sensitivities for singleplex and multiplex assays, respectively: for BuV ~4 copies/µl and ~9 copies/µl, for CuV ~5 copies/µl and ~6 copies/µl, and for TuV ~10 copies/µl and ~5 copies/µl, (Supplementary Fig. S1 in Study IV).

In the specificity testing, no amplification was seen of human DNA from several cultured cell lines or from high copy number plasmids of other human parvo- or polyomaviruses. Furthermore, no cross-amplification from high copy number plasmids of BuV1-3, TuV or CuV plasmids occurred.

CuV DNA in skin biopsies. CuV DNA was detected in the skin samples of 4/25 (16.0%) CTCL patients, and 4/136 (2.9%) transplant patients, whereas all 159 skin biopsies from 98 healthy adults were negative for CuV DNA. The difference of CuV DNA prevalence between the CTCL and transplant patients as well as between the CTCL and healthy adults, were statistically significant (P=0.00575 and P<0.0001, respectively, Pearson's χ^2 test).

In the CTCL cohort, all studied biopsies were from the cutaneous lymphoma lesion of the patients. To study the presence of CuV DNA in other tissues, additional samples were retrospectively obtained from the pathology archives from two CuV DNA-positive CTCL patients. From one patient, C-15, two prostate samples were available, and from the other, C-53, two lymph node, two prostate and eight skin samples, the latter including both non-malignant and malignant skin (pre-CTCL, CTCL and melanoma), were available for analysis. While the four prostate samples were CuV DNA-negative, CuV DNA was present in all skin and lymph node biopsies of patient C-53. Of note, the other CuV DNA-positive lymph node was sentinel to the prostate that did not harbor CuV DNA.

The highest CuV DNA quantities were observed in the CTCL lesions and in early CTCL tissue (lymphocytic atypia) with the viral load of 3.4E+07 to 7.3E+08 (per 1E+06 cells) (Table 2 in Study IV). The two melanoma biopsies had the lowest viral loads, 2.3E+02 and 7.2E+02 (per 1E+06 cells). The different CuV DNA-positive biopsies from the patient C-53 had been taken between 2012 and 2015 demonstrating that the patient harbored CuV DNA in his tissues at least for four years.

In the transplant cohort, CuV DNA was detected in four individuals (P-022, P-065, P-108 and P-A). Patients P-022 and P-108 were CuV-DNA positive in healthy skin, and these patients did not suffer from any skin malignancies (Table 9 and Table 2 in Study IV). Patients P-065 and P-A had skin cancer: basal cell carcinoma and squamous cell carcinoma, respectively. Both patients were CuV-DNA positive in all biopsies tested, and the viral load was higher in the cancerous tissues compared to the healthy skin (Table 9 and Table 2 in Study IV). Overall, when the CuV DNA prevalence in malignant skin biopsies (14.0%, 6/44) was compared with that in non-malignant skin biopsies (0.9%, 2/215), the difference reached statistical significance (P <0 .0001, Pearson's χ^2 test).

Table 9. ***CuV DNA and IgG results from transplant patients with CuV DNA detected in skin.*** *Adapted and modified from Study IV with permission.*

Patient	Tx date	Sampling date	Sampling site	Mean CuV DNA quantity*	CuV IgG OD value	Transplant and other info
P-022	5.1.2003	5.1.2003	Serum	NA	1.448	Liver tx
		28.1.2003	Serum	NA	0.554	CuV IgG pos prior to tx, and seropos for >10 years
		11.1.2013	**Healthy skin**	**4.51E+03**	-	
		11.1.2013	Serum	Negative	2.165	
P-065	2.12.1991	1.12.1991	Serum	NA	3.135	Liver tx
		30.12.1991	Serum	NA	0.805	CuV IgG pos prior to tx, and seropos for >21 years
		29.10.2013	**Healthy skin**	**3.92E+01**	-	
		29.10.2013	**Basal cell carcinoma**	**3.35E+03**	-	Diseased in 2015
		29.10.2013	Serum	Negative	1.477	
P-108	2.10.2009	30.6.2009	Serum	NA	1.978	Liver tx
		16.12.2014	**Healthy skin**	**2.71E+03**	-	CuV IgG pos prior to tx, and seropos for >5.5 years
		30.1.2015	Serum	Negative	1.241	
P-A	26.8.1986 and 25.6.1997	Unknown	**Healthy skin**	**2.54E+02**	-	Kidney tx at both times No serum available
		Unknown	**Pre-malignant (carcinoma in situ)**	**2.28E+03**	-	
		Unknown	**Squamous cell carcinoma**	**1.00E+04**	-	

*, per 1E+06 cells

NA, not analyzed; tx, transplant

To date, nine studies have reported analyses of CuV DNA, and in six studies CuV DNA was detected (including Study IV, Table 10). Of note, the NGS results of a CuV DNA-positive Danish melanoma patient have been published twice, however, they are counted here as one[139,192]. The presence of CuV DNA in skin samples has been reported in five studies and the CuV prevalence varied from 0% to 17% with the highest prevalence in cancer biopsies from CTCL patients (16%, Study IV) and in skin swabs from HIV-positive subjects (17%)[142] (Table 10).

In addition to Study IV and the study describing the discovery of CuV[36], Kreuter et al. found CuV DNA in CTCL patients from Germany: CuV DNA was detected exclusively in MF patients (6/71, 8.5%), whereas all patients (n=59) with other types of primary cutaneous B- or T-cell lymphomas were negative, including 20 patients with MF variants (folliculotropic MF, pagetoid reticulosis or Sézary syndrome)[141].

The CuV prevalence in the German study was lower (8.5%) than in Study IV (16%)[141]. Although there can be various explanations for the difference, one possibility could derive from the used methodology. The German study reported that all six CuV DNA-positive samples harbored 2.6-170 CuV DNA copies per cell, which is a relatively high load. In comparison, in Study IV only one CTCL sample had more than 1 CuV copy per cell (33.5 copies per cell in C-53), whereas the remaining three CuV DNA-positive samples had low

loads, 0.00003-0.01 per cell. The performance of the qPCR by Kreuter et al. was not reported[141], so its sensitivity and specificity could be lower than that in Study IV. The analysis of the primers used by Kreuter et al. showed that the forward primer had mismatches (0-2 nt) to the CuV sequences available in GenBank, which may cause lower specificity or sensitivity of the assay (Fig. 18). In addition, as both studies used FFPE samples and the FFPE procedure may degrade DNA, the differences in the fixation process may lead to different DNA-detection rates.

Samples from CTCL patients have been analyzed by Bergallo et al. with PCR and samples specifically from MF by Mollerup et al. with NGS[143,192]. Both, however, reported negative results, but the negativity may again derive from less sensitive methods or from the small sample size (n=11) in the latter study (Table 10).

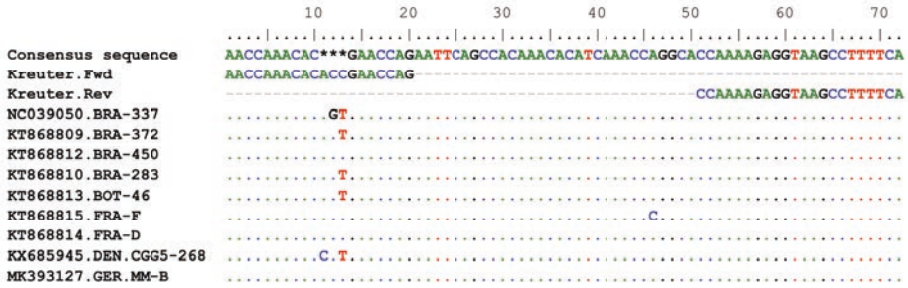


Figure 18. CuV primers used by Kreuter et al.[141] aligned with published CuV sequences. The reverse primer is the reverse complement sequence of the primer. The three stars in the consensus sequence depicts the mismatch locations in the primer. The probe is not shown as the sequence was not published (locked nucleic acid probe No. 5 produced by Roche). Dot (.), identical nt compared to the consensus sequence.

In Study IV, CuV DNA was detected in various skin biopsies in addition to the CTCL biopsies. In two transplant patients, CuV DNA was present both in healthy and in cancerous skin (basal cell or squamous cell carcinoma, Table 9). Furthermore, the CTCL patient C-53 harbored CuV DNA also in melanoma, atypical moles and dermatofibroma (benign), as well as in lymph nodes (Table 2 in Study IV). In line with Study IV, CuV has been detected in melanoma biopsies in two other studies[139,142]. In contrast, all skin samples from the healthy individuals in Study IV were CuV-DNA negative (n=98), but 3.8% (9/237) of the skin swabs from healthy individuals in Germany had CuV DNA. Remarkably, in the same German study a CuV DNA prevalence of 17.1% (35/205) was observed in skin swabs from HIV-positive individuals, which is also the highest CuV DNA prevalence in any study. However, it is not known whether CuV is more prevalent among the German HIV-positive patients in general or was the different sample type or sampling place causing the higher CuV DNA prevalence.

In all, the results indicate that CuV DNA can be detected in various skin specimens, including healthy and malignant skin as well as in lymph nodes, nevertheless, the role of CuV in any cancer or other diseases needs further studies.

Table 10. ***Studies describing analysis of CuV DNA.*** *The table includes all studies (Jan 2020) reporting the use on CuV PCR, both with and without CuV DNA findings, as well as all studies reporting NGS reads of CuV. The table is presented in chronological order.*

Study	Sample type	Country	Age (range)	Sampling time	n	Positive (%)	Symptoms	Detection method	Other info
Phan 2016[36]	**Feces**	**Brazil**	"Children"	Unknown	**245**	**4 (1.6%)**	GE	NGS, nPCR	Other pathogens in CuV pos samples: 1 sample with RVA and AAV*; 1 with AstV and AdV*; 1 with RVA and anellovirus*; 1 neg
	Feces	**Botswana**	"Children"	Unknown	**100**	**1 (1.0%)**	GE	nPCR	1 sample with PBV and anellovirus*
	Skin	**France**	"Adults"	Unknown	**NA**	**2**	CTCL	NGS, nPCR	ISH: Rare but strongly positive cells detected in 2/4 CTCL tissues positive by CuV PCR
	Skin	**France**	"Adults"	Unknown	**15**	**2 (13.3%)**	CTCL	nPCR	
	Skin	**France**	"Adults"	Unknown	**10**	**0**	Skin carcinoma	nPCR	
	Skin	**France**	"Adults"	Unknown	**NA**	**0**	Parapsoriasis	nPCR	
	Skin	**2 France; 6 unknown**	"Adults"	Unknown	**8**	**0**	Parapsoriasis	nPCR	
	Skin	**unknown**	"Adults"	Unknown	**8**	**0**	Eczema	nPCR	
	Skin	**unknown**	"Adults"	Unknown	**3**	**0**	Healthy	nPCR	
Mollerup 2017[139] Mollerup 2019[192]	**Skin**	**Denmark**	"Adults"	Unknown	**10#**	**1 (10%)**	Melanoma	NGS and real-time PCR	The 10 melanoma samples were part of larger NGS study: no CuV reads in other samples (n=197), including e.g. **11 MF**, 11 basal cell carcinoma, 10 oral c., 7 bladder c., 16 colon c., 20 breast c., 20 testicular c. and 47 various leukemia samples[192]
Kreuter 2017[193]	**Skin**	**Germany**	75 y	Unknown	**1**	**0**	C-ALCL	PCR	A case report. CuV PCR method not described, the sensitivity and specificity of the assay unknown.
Kreuter 2018[141]	**Lesional skin, FFPE**	**Germany**	Unknown	Unknown	**71**	**6 (8.5%)**	CTCL, MF	qPCR	CuV load: 1.3 - 85.0 copies per β-globin gene copy, median 19.4.
	Lesional skin, FFPE	**Germany**	Unknown	Unknown	**46**	**0**	Other	qPCR	Samples from other cutaneous T-cell and NK cell lymphomas and MF variants
	Lesional skin, FFPE	**Germany**	Unknown	Unknown	**13**	**0**	CBCL	qPCR	
Siqueira 2018[135]	**Feces**	**Venezuela, 3 isolated villages**	Mean 3 or 4 y§	2015	**60 (20 samples per village)**	**2 (3.3%)**	Healthy	NGS	NGS reads of CuV in two children from two isolated villages. Also BuV3 reads detected in the same villages, see Table 3
		Venezuela, Caracas	Mean 4 y	2015	**20**	**0**	Healthy	NGS	Urban city

Study	Sample type	Country	Age (range)	Sampling time	n	Positive (%)	Symptoms	Detection method	Other info
Bergallo 2018[143]	**CTCL skin**	**Italy**	Unknown	Unknown	**55**	**0**	CTCL	qPCR	The sensitivity and specificity of the qPCR unknown.
Dickinson 2019[164]	**JNA tissue, fresh**	**Finland**	Mean 22.7 y (17-33 y)	Unknown	**7**	**0**	JNA	Multiplex qPCR	Neg for BuV, TuV, CuV, B19V, HBoV1-4 and 13 HPyVs, expect one sample had MCPyV DNA
	Cancer and healthy tissues, fresh	**Finland**	Mean 61.2 y (31-86 y)	Unknown	**10 cancer, 8 healthy**	**0**	OPSCC/OSCC	Multiplex qPCR	Neg for BuV, TuV, CuV, and HBoV1-4. One patient had MCPyV, others HPyV neg. B19V in 7/10 patients
Study IV	**Skin biopsies, FFPE**	**Finland**	Median 64 y (32-89 y)	Unknown	**25**	**4 (16%)**	CTCL	Multiplex qPCR	CuV load: CTCL/early CTCL: 2.6E+01- 7.3E+08; other tissues from two patients with CuV DNA in the CTCL biopsy: neg - 5.5E+06 (per 1E+06 cells). Neg for BuV and TuV
	Skin biopsies, fresh	**Finland**	Median 62 y (22-83 y)	Unknown	**136**	**4 (2.9%)**	Transplant	Multiplex qPCR	CuV load: carcinoma/pre-carcinoma: 2.3E+03 - 1E+04; healthy skin: 3.9E+01- 4.5E+03 (per 1E+06 cells). Neg for BuV and TuV
	Skin biopsies, fresh	**Finland**	Median 43 y (18-67 y)	Unknown	**98**	**0**	Healthy	Multiplex qPCR	Neg for BuV and TuV
Wieland 2019[142]	**Melanoma tumor, FFPE**	**Germany**	Mean 64.8 y (21-93 y)	2002 - 2017	**179**	**2 (1.1%)**	Malignant melanoma	qPCR, real-time PCR	ISH: CuV detected only in the superficial layers and on the surface of the skin, not in the tumor cells. CuV load: 0.3 and 2.8 copies per β-globin gene copy.
	Melanoma metastases, FFPE	**Germany**	Unknown	Unknown	**52**	**0**	Malignant melanoma	qPCR	
	Skin swab, fresh	**Germany**	Mean 43 y (21-75 y)¤	Feb 2009 – Apr 2010¤	**205**	**35 (17.1%)**	HIV-positive	qPCR	CuV load in forehead skin swabs 0.004–268.75 DNA copies per β-globin gene copy
	Skin swab, fresh	**Germany**	Mean 47 y (18-79 y)¤	Feb 2009 – Apr 2010¤	**237**	**9 (3.8%)**	Healthy	qPCR	

*, NGS reads; #, the original article presented NGS and real-time PCR analysis of 10 melanoma patients[139]; §, depending on the village; ¤, age and sampling time data was derived from a previous study using the same cohorts with 7 additional individuals (5 HIV-positive patients and 2 healthy adults, total n=210 and n=239, respectively)[194]

AAV, adeno-associated virus; AdV, adenovirus; AstV, astrovirus; c., cancer; C-ALCL, CD30-positive primary cutaneous anaplastic large cell lymphoma; CBCL, cutaneous B-cell lymphoma; CTCL, cutaneous T-cell lymphoma; FFPE, formalin-fixed, paraffin-embedded; GE, gastroenteritis; HBoV, human bocavirus; HPyVs, human polyomaviruses; ISH, in situ hybridization; JNA, Juvenile nasopharyngeal angiofibroma; MCPyV, merkel cell polyomavirus; MF, mycosis fungoides; neg, negative; nPCR, nested PCR; OPSCC/OSCC, oropharyngeal and oral cavity squamous cell carcinoma; PBV, human picobirnavirus; qPCR, quantitative PCR; RVA, rotavirus A; y, year

Amplification of longer CuV fragment and CuV sequence analysis. All CuV qPCR positive samples were re-amplified with another PCR yielding 583 nt amplicons that included also the entire qPCR area. The longer amplicon was obtained from 6/7 of the CuV qPCR-positive fresh tissues from transplant patients, but only from 5/14 of the FFPE samples from the CTCL patients. It is known that formalin fixing degrades DNA, and it is often difficult or sometimes even impossible to amplify longer DNA fragments from FFPE tissues. Here all FFPE tissues with amplifiable 583 nt CuV DNA had very high CuV titres ranging from 4.8E+06 to 7.3E+08 (per 1E+06 cells), whereas FFPE tissues with CuV quantities of 2.3E+02 to 6.9E+05 (per 1E+06 cells) remained negative. In contrast, from the fresh tissues the 583-nt amplicon was obtained from a tissue with a CuV quantity as low as 3.9E+01 (per 1E+06 cells). In other words, viral load of one virus copy per cell was required for successful amplification from the FFPE tissues, but the load of several logs less was enough with the fresh tissues.

For the phylogenetic analysis the primer sites were removed, and the remaining 538 nt CuV sequence was used for analysis. The CuV sequences amplified from several tissues of one individual were shown to be identical, regardless whether the CuV was derived from healthy or (pre)malignant tissue. The result was the same in each of the three patients with multiple biopsies (P-065, P-A and C-53). In contrast to the identical sequences within each individual, the CuV sequences from different patients varied. No phylogenetic clustering, based on the sample type (skin or stool) or the geographical location of the patient harboring the virus, were seen when all available CuV sequences in GenBank were analyzed with the new sequences from this study (Fig. 1 in Study IV).

The identical sequence observed in each tissue from an individual could suggest that i) either there had not been any productive replication of the viral DNA in these tissues and the virus had persisted since the primary infection, ii) or that virus persisted elsewhere in the body, then replicated, and finally migrated to the skin only recently. It needs to be kept in mind that only a 538-nt piece of the VP2 region was used for this sequence comparison. A (nearly) full length genome of CuV should be amplified from each tissue to confirm whether there are mutations in the genome that could indicate active replication. However, with patient P-065, the serum sample corresponding to the skin biopsy was PCR negative demonstrating that the patient was not viremic at the sampling time.

BuV and TuV DNA in skin biopsies and the functionality of the multiplex qPCR assay. BuV and TuV DNA was not detected in the skin samples of any cohort. Although the lack of viral DNA in skin suggests that BuV and TuV are not persisting in skin tissue after primary infection like B19V or CuV, the study should be repeated with cohorts from endemic areas such as the Middle East or Africa (in the case of BuV). Among the studied Finnish subjects there were only few BuV1-3 or TuV IgG positives (see next chapter and Fig. 19), and thus the number of patients potentially harboring the viral DNA would be very low. Therefore, further studies are needed to confirm whether BuV or TuV can persist in skin tissue after primary infection.

As none of the tested skin biopsies had BuV or TuV DNA, the functionality of the assay amplifying the native single-stranded viral DNA from human samples was validated with fecal supernatants known to be BuV or TuV positive (the nearly full-length viral genomes[34,35] and the VP2 genes for creating VLPs had been previously amplified from these samples). By the novel multiplex qPCR, the BuV DNA quantities were very similar to the quantities detected with the BuV singleplex qPCR in 2012 (Table 11). The multiplex qPCR was also able to amplify TuV DNA from the TuV DNA-positive fecal supernatant. These results demonstrated that the multiplex qPCR could amplify the native ssDNA of BuV and TuV in human samples, and hence, the negative results from the skin biopsies were most likely caused by the true negativity of the samples instead of inadequate assay.

Table 11. ***Quantitative PCR results of BuV and TuV DNA-positive human fecal samples.*** *Reproduced from the supplementary material of Study IV with permission.*

Analysis year, method		2012, Singleplex qPCR	2017, Multiplex qPCR
Virus	Sample	copies/ml	copies/ml
BuV1	BF7	4.94E+03	1.72E+03
BuV1	BF86	8.12E+03	1.93E+03
BuV1	BF96	4.95E+04	1.38E+04
BuV2	BF39	3.74E+03	2.90E+03
TuV	Tu491	ND	4.22E+04

ND, not done; ml, milliliter of fecal supernatant.

BuV and CuV IgG prevalence in CTCL and transplant patients and in healthy adults. CuV IgG was the most prevalent in each cohort, whereas BuV1-3 IgG findings were less common or absent (Fig. 19). The CuV IgG prevalence was highest in the CTCL cohort (9.5%, 4/42), followed by the transplant cohort (6.5%, 8/124) and the healthy adults (3.8%, 3/78). Among the CTCL patients two patients had BuV1 IgG (4.8%, 2/42), while among the transplant recipients one patient had BuV1 IgG, one BuV2 IgG and one BuV3 IgG making the overall BuV-IgG prevalence of the cohort 2.4% (3/124) (Fig. 19).

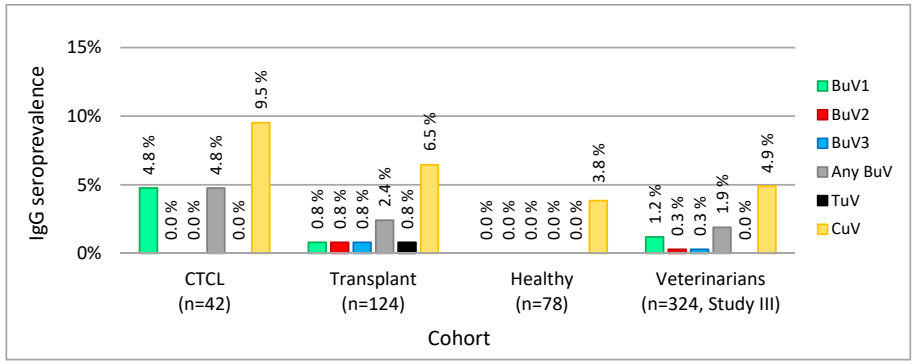


Figure 19. **IgG seroprevalence of BuV1-3, TuV and CuV in different adult cohorts in Finland.** "Veterinarians" is the same cohort than "Finland" in figure 16. The "any BuV" bar indicates the persons harboring BuV IgG towards one or more BuV types.

Nearly all seropositive individuals in these three cohorts had IgG towards a single protoparvovirus or virus type; only one transplant patient had IgG towards both BuV1 and CuV. None of the seroprevalence differences between the three cohorts were statistically significant. The IgG prevalences were also similar to the prevalence observed among veterinarians in Finland (4.9% for CuV and 1.9% for BuV) in Study III (Fig. 19).

Longevity and fluctuations of BuV and CuV IgG in immunosuppressed patients. From the transplant patients with BuV, TuV and/or CuV IgG detected in the original serum sample (n=10), additional serum samples were obtained from the archives and analyzed for protoparvovirus IgG. Archived serum samples were obtained also from the transplant patient P-108, with CuV DNA-positive skin, but who lacked the initial corresponding serum sample. Of these 11 patients (Table 12 and Table 4 in study IV), eight had the first available sample taken before or on the day of the organ transplant surgery, demonstrating the serological status of the patient prior to immunosuppression.

The EIA showed that five out of seven initially CuV IgG-positive subjects as well as the CuV DNA-positive patient P-108 harbored CuV IgG already before or on the day of organ transplantation (Table 12 and Table 4 in study IV). In addition, these patients remained CuV IgG positive for 5 to 21 years despite moderate immunosuppression (Table 12). As most of these patients were CuV IgG seropositive already in samples collected in 1991 or in 2001-2003, shows that these individuals had been infected with CuV in the early 2000, in 1990s or even before. This indicates that CuV has been circulating for a long time although it was described for the first time in 2016. Whether the individuals were infected with CuV in Finland or elsewhere, for example when travelling, cannot be defined based on the current data.

One patient (P-093) seroconverted to CuV during immunosuppression (Table 12). From patient P-060 a pre-transplantation serum sample was not available, however, samples taken 8 and 19 years after the organ transplantation were both CuV IgG positive. In patient P-055, who was TuV-IgG positive, the sample collected three days post-transplantation showed CuV IgG with a low OD. However, the follow-up samples from this patient were CuV IgG negative, and thus it can be hypothesized that this transient positivity could have been passively derived from blood products instead from a true infection. The TuV IgG positivity of this patient is discussed later in detail.

BuV1, BuV2 and BuV3 showed similar results to CuV: the BuV IgG-positive patients were BuV-IgG positive already before the transplantation and they remained seropositive for up to 11 years (Table 13 and Table 4 in study IV). The detection of long-lived BuV IgG is in line with our previous results that IgG can be detected in an individual for years (see chapter 4.2, results from Study II). Of note, the BuV1 IgG-positive patient was also CuV-IgG positive, and with both viruses the IgG was detected on the day of the transplantation as well as 9 years later, although with lower OD values.

Table 12. ***Longevity of CuV IgG in the transplant-patient cohort.*** *The IgG EIA results from follow-up samples taken at variable time points from the CuV sero- and/or genopositive patients. All positive results were confirmed with the competition assay. Negative IgG results (OD values) are illustrated with italics. Reproduced and modified from Study IV with permission.*

Patient	Transplant date	Sampling date	Age during sampling	CuV IgG OD value	Longevity of IgG
P-022	5.1.2003	5.1.2003	63 y	1.448	
		28.1.2003	63 y	0.554	
		11.1.2013*	73 y	2.165	10 y
P-065	2.12.1991	1.12.1991	47 y	3.135	
		30.12.1991	47 y	0.805	
		29.10.2013*	68 y	1.477	almost 22 y
P-108#	2.10.2009	30.6.2009	49 y	1.978	
		30.1.2015	55 y	1.241	5.5 y
P-012	14.11.2002	3.11.2002	40 y	0.415	
		1.12.2002	40 y	0.336	
		3.12.2012	50 y	0.117	10 y
P-037	7.3.2001	7.3.2001	42 y	1.256	
		27.3.2001	42 y	0. 830	
		13.3.2013	54 y	0.829	12 y
P-018§	21.11.2003	21.11.2003	44 y	0.352	
		31.12.2012	53 y	0.155	9 y
P-060	10.4.1994	27.11.2002	61 y	0.815	
		9.10.2013	72 y	0.660	almost 11 y
P-093	1st tx 19.3.1996; 2nd tx 26.8.1997	14.4.1997	58 y	*0.028*	The patient had seroconverted during follow up
		25.8.1998	60 y	*0.044*	
		23.9.1998	60 y	*0.030*	
		12.3.2014	75 y	0.294	
P-055	12.8.2001	15.8.2001	36 y	0.232¤	
See table 14		28.2.2003	37 y	0.019	
		24.9.2013	48 y	0.030	

*, CuV DNA was detected in the skin biopsy taken at same time than serum sample

#, the patient had CuV DNA in skin biopsy taken 16.12.2014, but no corresponding serum was available

§, the patient was also BuV1 IgG positive in both samples (see Table 13)

¤, IgG derived most likely from the blood transfusions/blood products, see text

Table 13. ***Longevity of BuV IgG in the transplant-patient cohort.*** *The IgG EIA column shows the OD values of the EIA for the virus mentioned in the first column. All positive results were confirmed with the competition assay. Negative IgG results (OD values) are illustrated with italics. Reproduced and modified from Study IV with permission.*

Virus	Patient	Transplant date	Sampling date	Age during sampling	IgG OD value	Longevity of IgG
BuV1	P-018*	21.11.2003	21.11.2003	44 y	2.066	
			31.12.2012	53 y	0.578	9 y
BuV2	P-034	11.3.2007	20.12.2006	58 y	0.169	
			21.4.2007	58 y	*0.085*	
			5.3.2013	64 y	0.168	6 y
BuV3	P-070	4.4.2002	23.2.2002	59 y	0.333	
			17.5.2002	59 y	0.120	
			12.11.2013	71 y	0.229	11.5 y

*, the patient was also CuV IgG positive in both samples (see Table 12)

In several cases a drop in IgG OD values was observed when comparing the sample taken before or at the time of transplantation with the sample taken 3-6 weeks after. Six such sample pairs were available (patients P-022, P-065, P-012, P-037 for CuV; P-034 for BuV2; and P-070 for BuV3, Tables 12 and 13), and in all the drop was apparent. The drop could be due to the immunosuppressive medication the patients receive after transplantation or to blood products. Because these virus infections are relatively infrequent in Finland, such drop after the transplantation and the recovery of the IgG level afterwards is easy to observe as the transfused blood is more likely to be protoparvovirus IgG negative and would not interfere with the results. In contrast, when analyzing the same samples for B19 IgG, a very prevalent parvovirus in Finland, the observed IgG fluctuations varied greatly (unpublished data).

TuV IgG prevalence in CTCL and transplant patients and in healthy adults, and longevity of the observed antibodies. Among the three cohorts containing 245 subjects with available serum samples, one individual harbored TuV IgG: transplant patient P-055 was TuV-IgG positive in two consecutive samples taken over 10 years apart (Table 14). Due to the rarity of the finding, the analysis was repeated with two different TuV antigen batches, nevertheless, the results were identical from both samples and indicated that this individual truly was TuV-IgG positive (Table 14).

In addition to the two serum samples with TuV IgG, a third serum sample was available from the same patient, however, no TuV IgG was detected (Table 14). This sample was taken three days post transplantation and 1.5 years before the first TuV IgG-positive serum sample. The negative result could be explained in two ways. First, it is possible that the patient was truly TuV-IgG negative at the time and seroconverted during the next 1.5 years. The second explanation could be that if the patient was TuV-IgG positive with a low titre before the surgery, the TuV IgG level three days post-surgery

Table 14. ***Competition assay results from TuV IgG-positive transplant patient P-055.*** *When the OD value without blocking is <0.1, the sample is considered negative (marked with *). The patient received a liver transplant 12.8.2001. The specific blocking is illustrated with bolded numbers.*

Sampling time		15.8.2001#			28.2.2003			24.9.2013		
Antigen in the well		**TuV, b1**	**TuV, b2**	**CuV**	**TuV, b1**	**TuV, b2**	**CuV**	**TuV, b1**	**TuV, b2**	**CuV**
Blocking antigen	none	0.055*	0.087*	0.218	0.355	0.462	0.019*	0.308	0.370	0.030*
	BuV1	0.068	0.060	0.234	0.331	0.443	ND	0.287	0.356	ND
	BuV2	0.055	0.064	0.217	0.177	0.244	ND	0.175	0.234	ND
	TuV, b1	0.072	0.054	ND	**0.052**	**0.020**	ND	**0.044**	**0.034**	ND
	TuV, b2	0.079	0.031	ND	**0.084**	**0.024**	ND	**0.066**	**0.034**	ND
	CuV	ND	ND	**0.045**	ND	ND	ND	ND	ND	ND
	H5	0.048	0.039	0.209	0.457	0.601	ND	0.378	0.505	ND

*, negative; #, the sample was taken three days post-transplant surgery during which the patient received blood transfusions. b1, TuV VLP batch 1; b2, TuV VLP batch 2; H5, HighFive insect cell lysate; ND, not done

could have been too low to be detected due to immunosuppressive medication. Such post-transplant drops in the IgG OD values of other viruses were observed in several transplant patients as discussed earlier. In addition, the TuV IgG negative sample contained CuV IgG, which was hypothesized to be derived from the blood transfusions as it was not detected in the follow up samples of P-055 (Tables 12 and 14). This transient positivity further shows that the antibodies detected in the sample most likely represented antibodies from the blood donor instead of the patient.

Nevertheless, to date, this is only the second time (a child in Study II being the first), when TuV IgG has been detected in an individual, and the first time ever, when longevity of TuV IgG has been documented.

CuV DNA and IgG correspondence. In this study, there were three patients (P-022, P-065 and P-108) harboring both CuV DNA and IgG. In each case, both the first serum sample, taken 5-21 years before the CuV DNA-positive skin biopsy, as well as the last serum sample, corresponding to the biopsy (or 6 weeks later in the case of P-108), were CuV-IgG positive (see also Fig. 21 for P-065). This suggests that CuV antibodies are long-lived, and that CuV DNA can persist for years, despite circulating antibodies, after acute infection similarly to that of B19V.

In addition to the above-mentioned three patients, there were ten patients with CuV IgG, but with no CuV DNA detected in the skin: two in the CTCL cohort, five in the transplant cohort and three in the healthy adults cohort. Thus, CuV DNA was detected in the skin of only 3/13 CuV IgG-positive subjects. With B19V, the correlation of sero- and genopositivity has been shown to be close to 100% when sensitive and specific IgG EIA and PCR assays are used[50,62]. So why were there many CuV DNA-negative individuals among the seropositives?

1. *Sampling place*. Already in 2016, it was shown by in situ hybridization that the CuV DNA-positive skin harbored very few, although strongly CuV positive cells[36]. In addition, a recent article described the presence of CuV in the superficial layer and surface of the skin, but not in deeper regions[142]. These findings point to that CuV is most likely not evenly distributed in the skin, and thus the sampling site might have a tremendous effect on CuV DNA detection and could explain some of the CuV DNA-negative findings in seropositive individuals.

2. *Sample type*. As CuV was discovered only recently, the permissive cell(s), symptoms during acute infection, the course of infection and development of immunity and persistence are not known. Although CuV DNA has been detected in skin of several individuals, the main place of persistence might be elsewhere in the body or for example in a migrating cell such as T cells or B cells. CuV could end up in skin only during certain conditions, for example in cancer, either as an infectious virus particle through the blood supply or as viral DNA inside a migrating cell.

3. *Variability of immune responses to acute infection*. The immune response towards CuV might vary between infected individuals, and it could be possible that in some

individuals, the immune system would not eradicate CuV after primary infection and therefore the virus persisted. Although such a phenomenon has not been observed among human parvoviruses, in other virus families this does play a role. For example, hepatitis C virus infection may be spontaneously cleared in some people, but others develop chronic infection.

4. *Used methods*. The CuV qPCR developed in Study IV was shown to be very sensitive and specific for CuV, nevertheless it is possible that samples with very low virus loads can remain negative. There can also be unknown CuV strains with mutations at the primer and probe sites, which can lead to no or reduced amplification. In this study, CuV sequences with a mismatch in the region of the 5' end of the forward primer were detected by the longer 583-nt PCR from several patients (Supplementary Fig. S2 in Study IV). Although this mismatch was located in the noncritical end of the primer, it demonstrates that there can be mutations in the amplified region.

CuV and BuV2 IgG cross-reactivity. In Study III, the interpretation of the competition EIA results between the cross-reactive CuV and BuV2 IgG was established. Analysis of samples from different cohorts demonstrated that the competition assay was able to distinguish cross-reactivity of these two viruses, exemplified by three samples from the Kenyan cohort presented in Table 7. Here, the original interpretation was further strengthened by the blocking results from the three patients harboring both CuV DNA and IgG: all competition assay results were identical to those in Study III. The CuV IgG reactivity was blocked completely with CuV and only lowered by BuV2, whereas BuV2 IgG reactivity was blocked completely both with BuV2 and CuV (Table 15). In addition, the identical competition assay results were obtained from all serum samples from these three CuV DNA- and IgG-positive individuals showing that the blocking characteristics and the antigen specificity could remain for over 21 years.

Table 15. ***BuV2 and CuV competition-assay results from the three patients harboring both CuV DNA and IgG.*** *In each sample the CuV IgG reactivity was blocked completely only with CuV, whereas the BuV2 IgG reactivity was blocked completely with both BuV2 and CuV. The numbers represent OD values from the IgG EIA, and blocking is illustrated with bolded numbers.*

Patient		P-022				P-065				P-108			
Sampling time		5.1.2003		11.1.2013*		1.12.1991		29.10.2013*		30.6.2009		30.1.2015#	
Antigen in the well		BuV2	CuV	BuV2	CuV	BuV2	CuV	BuV2	CuV	BuV2	CuV	BuV2	CuV
Blocking antigen	none	0.799	1.546	1.166	1.970	0.880	2.458	0.317	1.206	1.410	2.000	0.529	1.142
	BuV1	0.816	1.504	1.143	1.864	0.900	2.442	0.325	1.139	1.393	1.962	0.494	1.158
	BuV2	**0.025**	0.590	**0.011**	0.833	**0.033**	1.651	**0.018**	0.477	**0.045**	0.929	**0.016**	0.597
	CuV	**0.026**	**0.018**	**0.014**	**0.012**	**0.026**	**0.027**	**0.019**	**0.017**	**0.043**	**0.031**	**0.016**	**0.014**

*, Corresponding sample to the CuV DNA-positive skin biopsy; #, sample taken six weeks after the CuV DNA-positive skin biopsy

5 CONCLUDING REMARKS AND FUTURE PROSPECTS

In the beginning of the 21st century, the development of next generation sequencing methods revolutionized the identification of novel viruses from various sample materials. By using NGS, the amount and quality of sequence data increased tremendously, and the decreasing price made large-scale screening possible. However, the detected sequence itself provides only limited information on the properties of the virus, and thus further studies are needed to transfer the data from the computer into biological functions.

In this thesis, the epidemiology and clinical picture of three novel human parvoviruses were studied. Methods for DNA detection and quantification as well as for IgG antibody detection were developed and samples from altogether eleven cohorts were analyzed. The sample material included fecal samples from >1500 patients, nasal swabs from 955 children, and tissue biopsies from 260 individuals, all analyzed by qPCR, as well as serum samples from nearly 1500 individuals ranging four continents, analyzed by IgG EIA.

One of the major findings of the thesis was the remarkable difference in BuV seroprevalence observed between different geographical regions: in the Middle East and Africa the BuV IgG antibodies were detected in 56-85% of the adult population whereas in Finland and in the USA the seroprevalence was very low, <4% (Fig. 20). This indicates that Iraq, Iran and Kenya are endemic areas for BuV infections. In addition, the IgG EIA results of individuals from India suggested another endemic BuV area as well with high seroprevalence, however, due to the low sample number (n=6, of which 3 had BuV1 IgG),

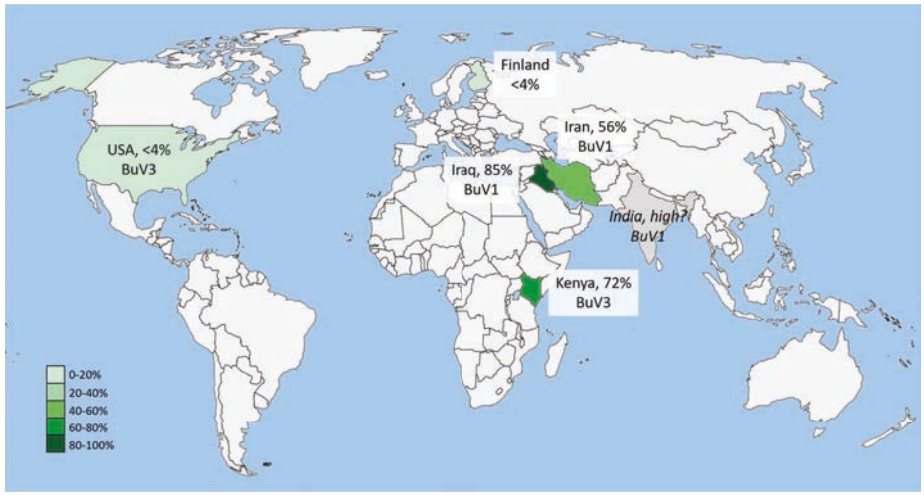


Figure 20. **Global BuV IgG seroprevalence among adults**. Combined seroprevalence of all BuV types is indicated as well as the predominant BuV type. In Finland none of the BuV types were very common, and in India, there were only 6 individuals tested (see text).

no clear conclusions could be made. The most prevalent BuV type varied as well, BuV1 being the most frequent in the Middle East (and India), whereas BuV3 dominated in Kenya.

According to the competition assay, the three BuV genotypes seem to represent serotypes, although low-level cross-reactions were observed within the high prevalence cohorts. Whether that is a common phenomenon in the endemic areas, in which the circulating BuVs may boost the immunity continuously or a result of subsequent infections of the different BuV types or even with currently unknown BuV types, the answer requires longitudinal follow up studies in the area.

By qPCR, we found BuV DNA in ten fecal samples from patients with gastroenteritis symptoms in Finland (0.6%, 10/1584), however, our findings did not strongly support the role of BuV in gastroenteritis. The results were in line with other publications, and the current consensus seems to be that BuV circulates worldwide, but it is detected only occasionally in fecal samples, and clusters of patients with BuV infection have not been observed. To elucidate the symptoms and clinical picture of the BuV infection, the next important step will be to identify patients with acute primary infection. As we have identified endemic areas with high BuV seroprevalence, patients from those regions would be a logical starting point when addressing this question. Another important question will be, do the three BuV types differ from each other clinically like HBoV1 differs from HBoV2-4 or are the infections very similar like with B19V genotypes.

Regarding TuV, there has not been any publications with positive DNA findings since the discovery, and in this thesis, we found TuV IgG in only one child and one transplant patient, and both with low OD values. Therefore, it is still unknown, whether TuV truly is a human virus or just accidentally infecting or occasionally detected in humans.

CuV, on the other hand, showed different characteristics than BuV or TuV: although the seroprevalence was low (0-9.5%) in every analyzed cohort, the CuV DNA was detected in skin biopsies of CTCL and transplant patients in Finland. The CuV DNA prevalence in the CTCL patients (16%) was significantly higher than in the transplant patients (2.9%) or in the healthy adults (0%). Furthermore, CuV IgG was detected in serum samples taken over 21 years before the CuV DNA-positive skin biopsy (Fig. 21), which suggested that CuV DNA can persist in the body for decades after primary infection similarly to that of B19V. However, we found CuV DNA in the skin of only 3 out of 13 CuV seropositive individuals, and thus it might be that either the skin cells are not the major site for the persistence, or CuV is not persisting in every individual.

In general, parvoviruses are not considered oncogenic, but instead some are described as oncotropic and oncolytic, both of which have been reported for animal protoparvoviruses. For example, the rat protoparvovirus H-1 has a natural tropism for human cancer cells and it is currently investigated as a therapeutic agent for glioblastoma (reviewed recently in [145]). However, only little is known of the molecular mechanisms and clinical impact of human parvovirus persistence in tissues. Are they inactive and dormant, or do they perhaps cause a long-term mild inflammation that could modify the host or immune cell activities. Previously, CuV has been detected by

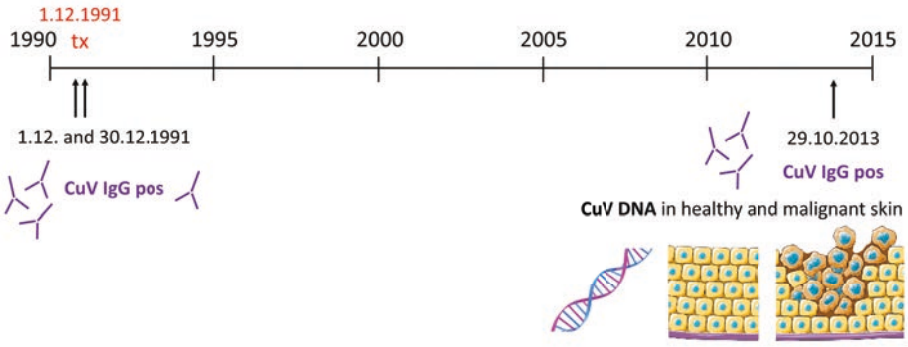


Figure 21. Schematic time-line illustration of the CuV IgG and DNA observed in patient P-065. CuV IgG was present already pre-tx with high OD (1.12.1991). One month post-tx a drop in the OD was observed, however, the IgG level recovered and was detected with high OD in 2013. CuV DNA was detected in both healthy and malignant skin in 2013. tx, liver transplantation

ISH only in a few, although strongly positive cells, in a CTCL biopsy[36] and only in the surface layer of a melanoma[142], which shows that the CuV is not present in every cancer cell. On the other hand, immunosuppression or other diseases may increase the risk of cancers associated with viruses, and thus the normally harmless virus persistence might turn into disease causing in the specific circumstances. Overall, even if some preliminary disease associations exist, the role of CuV in CTCL or other skin cancers needs further investigations.

6 ACKNOWLEDGEMENTS

This thesis work was carried out at the Department of Virology, Medicum, University of Helsinki. I warmly thank the head of the department, Professor Kalle Saksela for providing excellent facilities for work and an inspiring scientific atmosphere.

I would like to express my sincere gratitude and appreciation to my supervisor, docent Maria Söderlund-Venermo, for the great opportunity and even greater patience throughout this project. You have taught me a lot about science and familiarized me with the international community of parvovirus researchers. The collaborations you have started have been of immense importance for this thesis.

I would also like to thank my former supervisor, Professor Klaus Hedman. Especially I thank you for the guidance in writing. You have showed me how to formulate the message in such way that it suits the occasion in the best possible way (even when the explanation needs to be understood by "mökin mummo").

Reviewers of this thesis, docent Carita Savolainen-Kopra and docent Sisko Tauriainen are greatly acknowledged for their critical review and valuable suggestions to improve this thesis. Professor Kristina Broliden is acknowledged for accepting the role of opponent in my thesis dissertation.

I wish to thank all the collaborators and co-authors. Especially I am grateful for Dr. Eric Delwart and his team. Without your substantial contributions to the development of NGS and metagenomics, both in the wet lab and in silico, as well as the broad sample screening you have done, the viruses studied in this thesis would not have been discovered. Your collaboration is highly valued. Furthermore, I have been privileged to obtain invaluable clinical materials from various locations around the world. Maija Lappalainen, Eveliina Tarkka, Minna Paloniemi, Timo Vesikari, Arun Kumar, Rauli Franssila, Kamruddin Ahmed, Paula Kinnunen, Pikka Jokelainen, Haider Al-Hello, Ali Barakat, Eric Delwart, Reza Sadeghi, Farid Jalilian, Amir Majlesi, Moses Masika, Dufton Mwaengo, Omu Anzala, Olli Vapalahti, Sari Koskenmies, Nanna Fyhrquist, Anne Keinonen, Heikki Mäkisalo, Liisa Väkevä, Sari Pitkänen and Annamari Ranki, I thank you all warmly. Without these samples this thesis project could not have been done. I also thank you for your knowledge and expertise, and all your valuable comments to the publications.

I am thankful to all of my past and former co-workers at our group: Lea, Laura, Päivi, Heidi, Rauli, Kalle, Hanna, Anne, Mira, Mari, Claudia, Tingting, Arun, Reza, Yilin, Benedict, Inka, Väinö, Xuemeng, Mafe, Lari, Anna, Visa, Juuso, Usha, Yu, Maija, Henri, Outi, Diogo, and Rajita. Thank you all for spending time in the lab, office, lunch and coffee, and discussing ideas, science and everyday life. Lea is thanked for all the ELISA teaching and knowledge you have shared with me. Päivi is especially thanked for all the help in the lab throughout the years as well as for your friendship. Mari, I thank you for all interesting discussions both in the lab and outside, the enjoyable congress trips and for your friendship.

Inka, Väinö, Usha and Yu, you all have provided me important help in the lab by pipetting hundreds of samples. I would also like to thank Usha for continuing the work after me. You have already provided a lot of new and interesting data. All the colleagues in Department of Virology as well as in the building with special emphasis to the laboratory engineers are thanked for the nice atmosphere, companionship and all the help you have provided.

I would like to express my gratitude to my parents Sirpa and Eino, and to my sister Maria and her family for all the love and support. Without my father's example, I probably would not have selected microbiology as my field. I deeply thank Ossian for your support and patience through this decade long PhD process. And my daughter Siiri, you bring continuous happiness and surprise me every day with your insightful observations and humor. Thank you!

The Academy of Finland (projects 1122539 and 1257964), Biomedicum Helsinki Foundation, Clinical Chemistry Foundation, Finnish Society for Study of Infectious Disease, Finnish-Norwegian Medical Foundation, Foundation for Research on Viral Diseases, Helsinki Biomedical Graduate School (currently the Doctoral Programme in Biomedicine), the Helsinki University Hospital Research and Education Fund, Instrumentarium Foundation, Jenny and Antti Wihuri Foundation, Liv och Hälsa, Orion Foundation, Oscar Öflund Foundation, Otto A. Malm Foudation, Paulo Foundation, Sigrid Juselius Foundation, University of Helsinki Research Funds are thanked for financial support.

Helsinki, January 2020

Elina

7 REFERENCES

1. Brailovsky, C. [Research on the rat K virus (Parvovirus ratti). I. A method of titration by plaques and its application to the study of the multiplication cycle of the virus]. *Ann. Inst. Pasteur (Paris).* **110,** 49–59 (1966).
2. Cotmore, S. F., Agbandje-McKenna, M., Canuti, M., Chiorini, J. A., Eis-Hübinger, A.-M., Hughes, J., Mietzsch, M., Modha, S., Ogliastro, M., Pénzes, J. J., Pintel, D. J., Qiu, J., Söderlund-Venermo, M., Tattersall, P., Tijssen, P. & ICTV Report Consortium. ICTV Virus Taxonomy Profile: Parvoviridae. *J. Gen. Virol.* **100,** 367–368 (2019).
3. Pénzes, J. J., de Souza, W. M., Agbandje-McKenna, M. & Gifford, R. J. An ancient lineage of highly divergent parvoviruses infects both vertebrate and invertebrate hosts. *Viruses* **11,** 525 (2019).
4. Sukhu, L., Fasina, O., Burger, L., Rai, A., Qiu, J. & Pintel, D. J. Characterization of the nonstructural proteins of the bocavirus minute virus of canines. *J. Virol.* **87,** 1098–1104 (2013).
5. Naumer, M., Sonntag, F., Schmidt, K., Nieto, K., Panke, C., Davey, N. E., Popa-Wagner, R. & Kleinschmidt, J. A. Properties of the adeno-associated virus assembly-activating protein. *J. Virol.* **86,** 13038–48 (2012).
6. Zádori, Z., Szelei, J. & Tijssen, P. SAT: a late NS protein of porcine parvovirus. *J. Virol.* **79,** 13129–38 (2005).
7. Cotmore, S. F., Sturzenbecker, L. J. & Tattersall, P. The autonomous parvovirus MVM encodes two nonstructural proteins in addition to its capsid polypeptides. *Virology* **129,** 333–343 (1983).
8. Naeger, L. K., Cater, J. & Pintel, D. J. The small nonstructural protein (NS2) of the parvovirus minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner. *J. Virol.* **64,** 6166–75 (1990).
9. Alexandersen, S., Bloom, M. E. & Perryman, S. Detailed transcription map of Aleutian mink disease parvovirus. *J. Virol.* **62,** 3684–94 (1988).
10. Cotmore, S. F. & Tattersall, P. Parvoviruses: small does not mean simple. *Annu. Rev. Virol.* **1,** 517–537 (2014).
11. Tattersall, P. & Ward, D. C. Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. *Nature* **263,** 106–109 (1976).
12. Cotmore, S. F., McKie, V. C., Anderson, L. J., Astell, C. R. & Tattersall, P. Identification of the major structural and nonstructural proteins encoded by human parvovirus B19 and mapping of their genes by procaryotic expression of isolated genomic fragments. *J. Virol.* **60,** 548–57 (1986).
13. Mietzsch, M., Pénzes, J. J., Agbandje-McKenna, M., Mietzsch, M., Pénzes, J. J. & Agbandje-McKenna, M. Twenty-five years of structural parvovirology. *Viruses* **11,** 362 (2019).
14. Kajigaya, S., Fujii, H., Field, A., Anderson, S., Rosenfeld, S., Anderson, L. J., Shimada, T. & Young, N. S. Self-assembled B19 parvovirus capsids, produced in a baculovirus system, are antigenically and immunogenically similar to native virions. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 4646–4650 (1991).
15. Bräuniger, S., Peters, J., Borchers, U., Kao, M. & Borchers, U. Further studies on thermal resistance of bovine parvovirus against moist and dry heat. *Int. J. Hyg. Environ. Health* **203,** 71–75 (2000).
16. Kaufmann, B., Simpson, A. A. & Rossmann, M. G. The structure of human parvovirus B19. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 11628–11633 (2004).

17. Gurda, B. L., Parent, K. N., Bladek, H., Sinkovits, R. S., DiMattia, M. A., Rence, C., Castro, A., McKenna, R., Olson, N., Brown, K., Baker, T. S. & Agbandje-McKenna, M. Human bocavirus capsid structure: Insights into the structural repertoire of the parvoviridae. *J. Virol.* **84,** 5880–5889 (2010).

18. Kilham, L. & Olivier, L. J. A latent virus of rats isolated in tissue culture. *Virology* **7,** 428–437 (1959).

19. Atchison, R. W., Casto, B. C. & Hammon, W. M. Adenovirus-associated defective virus particles. *Science* **149,** 754–755 (1965).

20. Cossart, Y. E., Cant, B., Field, A. M. & Widdows, D. Parvovirus-like particles in human sera. *Lancet* **305,** 72–73 (1975).

21. Söderlund-Venermo, M. Emerging human parvoviruses: The rocky road to fame. *Annu. Rev. Virol.* **6,** 71–91 (2019).

22. Allander, T., Emerson, S. U., Engle, R. E., Purcell, R. H. & Bukh, J. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 11609 (2001).

23. Allander, T., Tammi, M. T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A. & Andersson, B. Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 12891–6 (2005).

24. Jones, M. S., Kapoor, A., Lukashov, V. V, Simmonds, P., Hecht, F. & Delwart, E. New DNA viruses identified in patients with acute viral infection syndrome. *J. Virol.* **79,** 8230–6 (2005).

25. Fryer, J. F., Delwart, E., Hecht, F. M., Bernardin, F., Jones, M. S., Shah, N. & Baylis, S. A. Frequent detection of the parvoviruses, PARV4 and PARV5, in plasma from blood donors and symptomatic individuals. *Transfusion* **47,** 1054–1061 (2007).

26. Simmonds, P., Douglas, J., Bestetti, G., Longhi, E., Antinori, S., Parravicini, C. & Corbellino, M. A third genotype of the human parvovirus PARV4 in sub-Saharan Africa. *J. Gen. Virol.* **89,** 2299–2302 (2008).

27. Kapoor, A., Slikas, E., Simmonds, P., Chieochansin, T., Naeem, A., Shaukat, S., Alam, M. M., Sharif, S., Angez, M., Zaidi, S. & Delwart, E. A newly identified bocavirus species in human stool. *J. Infect. Dis.* **199,** 196–200 (2009).

28. Arthur, J. L., Higgins, G. D., Davidson, G. P., Givney, R. C. & Ratcliff, R. M. A novel bocavirus associated with acute gastroenteritis in Australian children. *PLoS Pathog.* **5,** e1000391 (2009).

29. Kapoor, A., Simmonds, P., Slikas, E., Li, L., Bodhidatta, L., Sethabutr, O., Triki, H., Bahri, O., Oderinde, B. S., Baba, M. M., Bukbuk, D. N., Besser, J., Bartkus, J. & Delwart, E. Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections. *J. Infect. Dis.* **201,** 1633–1643 (2010).

30. Väisänen, E., Lahtinen, A., Eis-Hübinger, A. M., Lappalainen, M., Hedman, K. & Söderlund-Venermo, M. A two-step real-time PCR assay for quantitation and genotyping of human parvovirus 4. *J. Virol. Methods* **195,** 106–111 (2014).

31. Matthews, P. C., Sharp, C., Simmonds, P. & Klenerman, P. Human parvovirus 4 'PARV4' remains elusive despite a decade of study. *F1000Research* **6,** 82 (2017).

32. Matthews, P. C., Malik, A., Simmons, R., Sharp, C., Simmonds, P. & Klenerman, P. PARV4: An emerging tetraparvovirus. *PLoS Pathog.* **10,** e1004036 (2014).

33. Cotmore, S. F., Agbandje-McKenna, M., Chiorini, J. A., Mukha, D. V, Pintel, D. J., Qiu, J., Söderlund-Venermo, M., Tattersall, P., Tijssen, P., Gatherer, D. & Davison, A. J. The family Parvoviridae. *Arch. Virol.* **159,** 1239–47 (2014).

34. Phan, T. G., Vo, N. P., Bonkoungou, I. J. O., Kapoor, A., Barro, N., O'Ryan, M., Kapusinszky, B., Wang, C. & Delwart, E. Acute diarrhea in West African children: diverse enteric viruses and a novel parvovirus genus. *J. Virol.* **86,** 11024–30 (2012).

35. Phan, T. G., Sdiri-Loulizi, K., Aouni, M., Ambert-Balay, K., Pothier, P., Deng, X. & Delwart, E. New parvovirus in child with unexplained diarrhea, Tunisia. *Emerg. Infect. Dis.* **20,** 1911–3 (2014).

36. Phan, T. G., Dreno, B., da Costa, A. C., Li, L., Orlandi, P., Deng, X., Kapusinszky, B., Siqueira, J., Knol, A.-C., Halary, F., Dantal, J., Alexander, K. A., Pesavento, P. A. & Delwart, E. A new protoparvovirus in human fecal samples and cutaneous T cell lymphomas (mycosis fungoides). *Virology* **496,** 299–305 (2016).

37. Deiss, V., Tratschin, J.-D., Weitz, M. & Siegl, G. Cloning of the human parvovirus B19 genome and structural analysis of its palindromic termini. *Virology* **175,** 247–254 (1990).

38. Cotmore, S. F. & Tattersall, P. Characterization and molecular cloning of a human parvovirus genome. *Science* **226,** 1161–5 (1984).

39. Qiu, J., Söderlund-Venermo, M. & Young, N. S. Human parvoviruses. *Clin. Microbiol. Rev.* **30,** 43–113 (2017).

40. Chen, A. Y., Zhang, E. Y., Guan, W., Cheng, F., Kleiboeker, S., Yankee, T. M. & Qiu, J. The small 11kDa nonstructural protein of human parvovirus B19 plays a key role in inducing apoptosis during B19 virus infection of primary erythroid progenitor cells. *Blood* **115,** 1070–1080 (2010).

41. Brown, K. E., Anderson, S. M. & Young, N. S. Erythrocyte P antigen: cellular receptor for B19 parvovirus. *Science* **262,** 114–7 (1993).

42. Weigel-Kelley, K. A., Yoder, M. C. & Srivastava, A. α5β1 integrin as a cellular coreceptor for human parvovirus B19: requirement of functional activation of β1 integrin for viral entry. *Blood* **102,** 3927–3933 (2003).

43. Munakata, Y., Saito-Ito, T., Kumura-Ishii, K., Huang, J., Kodera, T., Ishii, T., Hirabayashi, Y., Koyanagi, Y. & Sasaki, T. Ku80 autoantigen as a cellular coreceptor for human parvovirus B19 infection. *Blood* **106,** 3449–3456 (2005).

44. Bieri, J. & Ros, C. Globoside is dispensable for parvovirus B19 entry but essential at a postentry step for productive infection. *J. Virol.* **93,** e00972-19 (2019).

45. Brown, K. E., Hibbs, J. R., Gallinella, G., Anderson, S. M., Lehman, E. D., McCarthy, P. & Young, N. S. Resistance to parvovirus B19 infection due to lack of virus receptor (erythrocyte P antigen). *N. Engl. J. Med.* **330,** 1192–1196 (1994).

46. Young, N. S., Mortimer, P. P., Moore, J. G. & Humphries, R. K. Characterization of a virus that causes transient aplastic crisis. *J. Clin. Invest.* **73,** 224–30 (1984).

47. Young, N., Harrison, M., Moore, J., Mortimer, P. & Humphries, R. K. Direct demonstration of the human parvovirus in erythroid progenitor cells infected in vitro. *J. Clin. Invest.* **74,** 2024–32 (1984).

48. Munakata, Y., Kato, I., Saito, T., Kodera, T., Ishii, K. K. & Sasaki, T. Human parvovirus B19 infection of monocytic cell line U937 and antibody-dependent enhancement. *Virology* **345,** 251–257 (2006).

49. von Kietzell, K., Pozzuto, T., Heilbronn, R., Grössl, T., Fechner, H. & Weger, S. Antibody-mediated enhancement of parvovirus B19 uptake into endothelial cells mediated by a receptor for complement factor C1q. *J. Virol.* **88,** 8102–15 (2014).

50. Pyöriä, L., Toppinen, M., Mäntylä, E., Hedman, L., Aaltonen, L.-M., Vihinen-Ranta, M., Ilmarinen, T., Söderlund-Venermo, M., Hedman, K. & Perdomo, M. F. Extinct type of human parvovirus B19 persists in tonsillar B cells. *Nat. Commun.* **8,** 14930 (2017).

51. Vuorinen, T., Lammintausta, K., Kotilainen, P. & Nikkari, S. Presence of parvovirus B19 DNA in chronic urticaric and healthy human skin. *J. Clin. Virol.* **25,** 217–221 (2002).

52. Söderlund, M., von Essen, R., Haapasaari, J., Kiistala, U., Kiviluoto, O. & Hedman, K. Persistence of parvovirus B19 DNA in synovial membranes of young patients with and without chronic arthropathy. *Lancet* **349,** 1063–1065 (1997).

53. Norja, P., Hokynar, K., Aaltonen, L.-M., Chen, R., Ranki, A., Partio, E. K., Kiviluoto, O., Davidkin, I., Leivo, T., Eis-Hübinger, A. M., Schneider, B., Fischer, H.-P., Tolba, R., Vapalahti, O., Vaheri, A., Söderlund-Venermo, M. & Hedman, K. Bioportfolio: lifelong persistence of variant and prototypic erythrovirus DNA genomes in human tissue. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 7450–3 (2006).

54. Servant, A., Laperche, S., Lallemand, F., Marinho, V., De Saint Maur, G., Meritet, J. F. & Garbarg-Chenon, A. Genetic diversity within human erythroviruses: identification of three genotypes. *J. Virol.* **76,** 9124–34 (2002).

55. Hokynar, K., Söderlund-Venermo, M., Pesonen, M., Ranki, A., Kiviluoto, O., Partio, E. K. & Hedman, K. A new parvovirus genotype persistent in human skin. *Virology* **302,** 224–228 (2002).

56. Hübschen, J. M., Mihneva, Z., Mentis, A. F., Schneider, F., Aboudy, Y., Grossman, Z., Rudich, H., Kasymbekova, K., Sarv, I., Nedeljkovic, J., Tahita, M. C., Tarnagda, Z., Ouedraogo, J.-B., Gerasimova, A. G., Moskaleva, T. N., Tikhonova, N. T., Chitadze, N., Forbi, J. C., Faneye, A. O., Otegbayo, J. A., Charpentier, E. & Muller, C. P. Phylogenetic analysis of human parvovirus B19 sequences from eleven different countries confirms the predominance of genotype 1 and suggests the spread of genotype 3b. *J. Clin. Microbiol.* **47,** 3735–8 (2009).

57. Candotti, D., Etiz, N., Parsyan, A. & Allain, J.-P. Identification and characterization of persistent human erythrovirus infection in blood donor samples. *J. Virol.* **78,** 12169–78 (2004).

58. Jain, P., Jain, A., Prakash, S., Khan, D. N., Singh, D. D., Kumar, A., Moulik, N. R. & Chandra, T. Prevalence and genotypic characterization of human parvovirus B19 in children with hemato-oncological disorders in North India. *J. Med. Virol.* **87,** 303–309 (2015).

59. Aydin, H., Timurkan, M. O. & Aktas, O. Genetic diversity of primate erythroparvovirus 1 between 2009 and 2016: First report from Turkey. *Microb. Pathog.* **129,** 131–135 (2019).

60. Nguyen, Q. T., Wong, S., Heegaard, E. D. & Brown, K. E. Identification and characterization of a second novel human erythrovirus variant, A6. *Virology* **301,** 374–380 (2002).

61. Sanabani, S., Neto, W. K., Pereira, J. & Sabino, E. C. Sequence variability of human erythroviruses present in bone marrow of Brazilian patients with various parvovirus B19-related hematological symptoms. *J. Clin. Microbiol.* **44,** 604–6 (2006).

62. Toppinen, M., Perdomo, M. F., Palo, J. U., Simmonds, P., Lycett, S. J., Söderlund-Venermo, M., Sajantila, A. & Hedman, K. Bones hold the key to DNA virus history and epidemiology. *Sci. Rep.* **5,** 17226 (2015).

63. Ekman, A., Hokynar, K., Kakkola, L., Kantola, K., Hedman, L., Bondén, H., Gessner, M., Aberham, C., Norja, P., Miettinen, S., Hedman, K. & Söderlund-Venermo, M. Biological and immunological relations among human parvovirus B19 genotypes 1 to 3. *J. Virol.* **81,** 6927–35 (2007).

64. Parsyan, A., Kerr, S., Owusu-Ofori, S., Elliott, G. & Allain, J.-P. Reactivity of genotype-specific recombinant proteins of human erythrovirus B19 with plasmas from areas where genotype 1 or 3 is endemic. *J. Clin. Microbiol.* **44,** 1367–75 (2006).

65. Heegaard, E. D., Qvortrup, K. & Christensen, J. Baculovirus expression of erythrovirus V9 capsids and screening by ELISA: Serologic cross-reactivity with erythrovirus B19. *J. Med. Virol.* **66,** 246–252 (2002).

66. Anderson, M. J., Higgins, P. G., Davis, L. R., Willman, J. S., Jones, S. E., Kidd, I. M., Pattison, J. R. & Tyrrell, D. A. J. Experimental parvoviral infection in humans. *J. Infect. Dis.* **152,** 257–265 (1985).

67. Anderson, M. J., Davis, L. R., Hodgson, J., Jones, S. E., Murtaza, L., Pattison, J. R., Stroud, C. E. & White, J. M. Occurrence of infection with a parvovirus-like agent in children with sickle cell anaemia during a two-year period. *J. Clin. Pathol.* **35,** 744–9 (1982).

68. Mortimer, P. P., Luban, N. L. C., Kelleher, J. F. & Cohen, B. J. Transmission of serum parvovirus-like virus by clotting-factor concentrates. *Lancet* **322,** 482–484 (1983).

69. Brown, T., Anand, A., Ritchie, L. D., Clewley, J. P. & Reid, T. M. S. Intrauterine parvovirus infection associated with hydrops fetalis. *Lancet* **324,** 1033–1034 (1984).

70. Anderson, M. J., Lewis, E., Kidd, I. M., Hall, S. M. & Cohen, B. J. An outbreak of erythema infectiosum associated with human parvovirus infection. *J. Hyg. (Lond).* **93,** 85–93 (1984).

71. Woolf, A. D., Campion, G. V., Chishick, A., Wise, S., Cohen, B. J., Klouda, P. T., Caul, O. & Dieppe, P. A. Clinical manifestations of human parvovirus B19 in adults. *Arch. Intern. Med.* **149,** 1153 (1989).

72. Kurtzman, G., Frickhofen, N., Kimball, J., Jenkins, D. W., Nienhuis, A. W. & Young, N. S. Pure red-cell aplasia of 10 years' duration due to persistent parvovirus B19 infection and its cure with immunoglobulin therapy. *N. Engl. J. Med.* **321,** 519–523 (1989).

73. Rao, S. P., Miller, S. T. & Cohen, B. J. Severe anemia due to B19 parvovirus infection in children with acute leukemia in remission. *Am. J. Pediatr. Hematol. Oncol.* **12,** 194–197 (1990).

74. Pattison, J. R., Jones, S. E., Hodgson, J., Davis, L. R., White, J. M., Stroud, C. E. & Murtaza, L. Parvovirus infections and hypoplastic crisis in sickle-cell anaemia. *Lancet* **317,** 664–665 (1981).

75. Serjeant, G. R., Mason, K., Topley, J. M., Serjeant, B. E., Pattison, J. R., Jones, S. E. & Mohamed, R. Outbreak of aplastic crises in sickle cell anaemia associated with parvovirus-like agent. *Lancet* **318,** 595–597 (1981).

76. Valeur-Jensen, A. K., Pedersen, C. B., Westergaard, T., Jensen, I. P., Lebech, M., Andersen, P. K., Aaby, P., Pedersen, B. N. & Melbye, M. Risk factors for parvovirus B19 infection in pregnancy. *JAMA* **281,** 1099 (1999).

77. Koch, W. C. & Adler, S. P. Human parvovirus B19 infections in women of childbearing age and within families. *Pediatr. Infect. Dis. J.* **8,** 83–7 (1989).

78. Riipinen, A., Sallmén, M., Hedman, L., Ojajärvi, A., Lindbohm, M.-L., Meriluoto, M., Surcel, H.-M., Taskinen, H., Nuutila, M., Karikoski, R., Hedman, K. & Söderlund-Venermo, M. Increased risk of human parvovirus B19 infection in day-care employees: a cohort study among pregnant workers during an epidemic in Finland. *Occup. Environ. Med.* **71,** 836–41 (2014).

79. Skjöldebrand-Sparre, L., Fridell, E., Nyman, M. & Wahren, B. A prospective study of antibodies against parvovirus B19 in pregnancy. *Acta Obstet. Gynecol. Scand.* **75,** 336–339 (1996).

80. Public Health Laboratory Service Working Party on Fifth Disease. Prospective study of human parvovirus (B19) infection in pregnancy. *BMJ* **300,** 1166–70 (1990).

81. Gratacos, E., Torres, P.-J., Vidal, J., Antolin, E., Costa, J., Jimenez de Anta, M. T., Cararach, V., Alonso, P. L. & Fortuny, A. The incidence of human parvovirus B19 infection during pregnancy and its impact on perinatal outcome. *J. Infect. Dis.* **171,** 1360–1363 (1995).

82. Rodis, J. F., Quinn, D. L., Gary, G. W., Anderson, L. J., Rosengren, S., Cartter, M. L., Campbell, W. A. & Vintzileos, A. M. Management and outcomes of pregnancies complicated by human B19 parvovirus infection: A prospective study. *Am. J. Obstet. Gynecol.* **163,** 1168–1171 (1990).

83. Lassen, J., Bager, P., Wohlfahrt, J., Bottiger, B. & Melbye, M. Parvovirus B19 infection in pregnancy and subsequent morbidity and mortality in offspring. *Int. J. Epidemiol.* **42,** 1070–1076 (2013).

84. De Jong, E. P., Lindenburg, I. T., van Klink, J. M., Oepkes, D., van Kamp, I. L., Walther, F. J. & Lopriore, E. Intrauterine transfusion for parvovirus B19 infection: long-term neurodevelopmental outcome. *Am. J. Obstet. Gynecol.* **206,** 204.e1-204.e5 (2012).

85. Böhmer, A., Schildgen, V., Lüsebrink, J., Ziegler, S., Tillmann, R. L., Kleines, M. & Schildgen, O. Novel application for isothermal nucleic acid sequence-based amplification (NASBA). *J. Virol. Methods* **158,** 199–201 (2009).

86. Dijkman, R., Koekkoek, S. M., Molenkamp, R., Schildgen, O. & van der Hoek, L. Human bocavirus can be cultured in differentiated human airway epithelial cells. *J. Virol.* **83,** 7739–48 (2009).

87. Chen, A. Y., Cheng, F., Lou, S., Luo, Y., Liu, Z., Delwart, E., Pintel, D. & Qiu, J. Characterization of the gene expression profile of human bocavirus. *Virology* **403,** 145–154 (2010).

88. Shen, W., Deng, X., Zou, W., Cheng, F., Engelhardt, J. F., Yan, Z. & Qiu, J. Identification and functional analysis of novel nonstructural proteins of human bocavirus 1. *J. Virol.* **89,** 10097–109 (2015).

89. Zou, W., Xiong, M., Deng, X., Engelhardt, J. F., Yan, Z. & Qiu, J. A comprehensive RNA-seq analysis of human bocavirus 1 transcripts in infected human airway epithelium. *Viruses* **11,** (2019).

90. Deng, X., Yan, Z., Cheng, F., Engelhardt, J. F. & Qiu, J. Replication of an autonomous human parvovirus in non-dividing human airway epithelium is facilitated through the DNA damage and repair pathways. *PLOS Pathog.* **12,** e1005399 (2016).

91. Lüsebrink, J., Schildgen, V., Tillmann, R. L., Wittleben, F., Böhmer, A., Müller, A. & Schildgen, O. Detection of head-to-tail DNA sequences of human bocavirus in clinical samples. *PLoS One* **6,** e19457 (2011).

92. Kapoor, A., Hornig, M., Asokan, A., Williams, B., Henriquez, J. A. & Lipkin, W. I. Bocavirus episome in infected human tissue contains non-identical termini. *PLoS One* **6,** e21362 (2011).

93. Zhao, H., Zhao, L., Sun, Y., Qian, Y., Liu, L., Jia, L., Zhang, Y. & Dong, H. Detection of a bocavirus circular genome in fecal specimens from children with acute diarrhea in Beijing, China. *PLoS One* **7,** e48980 (2012).

94. Babkin, I. V., Tyumentsev, A. I., Tikunov, A. Y., Zhirakovskaia, E. V., Netesov, S. V. & Tikunova, N. V. A study of the human bocavirus replicative genome structures. *Virus Res.* **195,** 196–202 (2015).

95. Christensen, A., Kesti, O., Elenius, V., Eskola, A. L., Døllner, H., Altunbulakli, C., Akdis, C. A., Söderlund-Venermo, M. & Jartti, T. Human bocaviruses and paediatric infections. *Lancet Child Adolesc. Heal.* **3,** 418–426 (2019).

96. Riipinen, A., Väisänen, E., Lahtinen, A., Karikoski, R., Nuutila, M., Surcel, H.-M., Taskinen, H., Hedman, K. & Söderlund-Venermo, M. Absence of human bocavirus from deceased fetuses and their mothers. *J. Clin. Virol.* **47,** (2010).

97. Enders, M., Lindner, J., Wenzel, J. J., Baisch, C., Schalasta, G., Enders, G. & Modrow, S. No detection of human bocavirus in amniotic fluid samples from fetuses with hydrops or isolated effusions. *J. Clin. Virol.* **45,** 300–303 (2009).

98. Hansen, M., Brockmann, M., Schildgen, V. & Schildgen, O. Human bocavirus is detected in human placenta and aborted tissues. *Influenza Other Respi. Viruses* **13,** 106–109 (2019).

99. Modrow, S., Wenzel, J. J., Schimanski, S., Schwarzbeck, J., Rothe, U., Oldenburg, J., Jilg, W. & Eis-Hübinger, A. M. Prevalence of nucleic acid sequences specific for human parvoviruses, hepatitis A and hepatitis E viruses in coagulation factor concentrates. *Vox Sang.* **100,** 351–358 (2011).

100. Li, H., He, M., Zeng, P., Gao, Z., Bian, G., Yang, C. & Li, W. The genomic and seroprevalence of human bocavirus in healthy Chinese plasma donors and plasma derivatives. *Transfusion* **55,** 154–163 (2015).

101. Martin, E. T., Kuypers, J., McRoberts, J. P., Englund, J. A. & Zerr, D. M. Human bocavirus 1 primary infection and shedding in infants. *J. Infect. Dis.* **212,** 516–24 (2015).

102. Söderlund-Venermo, M., Lahtinen, A., Jartti, T., Hedman, L., Kemppainen, K., Lehtinen, P., Allander, T., Ruuskanen, O. & Hedman, K. Clinical assessment and improved diagnosis of bocavirus-induced wheezing in children, Finland. *Emerg. Infect. Dis.* **15,** 1423 (2009).

103. Kantola, K., Hedman, L., Tanner, L., Simell, V., Makinen, M., Partanen, J., Sadeghi, M., Veijola, R., Knip, M., Ilonen, J., Hyoty, H., Toppari, J., Simell, O., Hedman, K. & Söderlund-Venermo, M. B-cell responses to human bocaviruses 1-4: New insights from a childhood follow-up study. *PLoS One* **10,** e0139096 (2015).

104. Hedman, L., Söderlund-Venermo, M., Jartti, T., Ruuskanen, O. & Hedman, K. Dating of human bocavirus infection with protein-denaturing IgG-avidity assays - Secondary immune activations are ubiquitous in immunocompetent adults. *J. Clin. Virol.* **48,** 44–48 (2010).

105. Kantola, K., Hedman, L., Arthur, J. L., Alibeto, A., Delwart, E., Jartti, T., Ruuskanen, O., Hedman, K. & Söderlund-Venermo, M. Seroepidemiology of human bocaviruses 1-4. *J. Infect. Dis.* **204,** 1403–1412 (2011).

106. Li, X., Kantola, K., Hedman, L., Arku, B., Hedman, K. & Söderlund-Venermo, M. Original antigenic sin with human bocaviruses 1–4. *J. Gen. Virol.* **96,** 3099–3108 (2015).

107. Endo, R., Ishiguro, N., Kikuta, H., Teramoto, S., Shirkoohi, R., Ma, X., Ebihara, T., Ishiko, H. & Ariga, T. Seroepidemiology of human bocavirus in Hokkaido prefecture, Japan. *J. Clin. Microbiol.* **45,** 3218–23 (2007).

108. Sadeghi, M., Kantola, K., Finnegan, D. P. J., McCaughey, C., Hedman, L., Söderlund-Venermo, M. & Hedman, K. Possible involvement of human bocavirus 1 in the death of a middle-aged immunosuppressed patient. *J. Clin. Microbiol.* **51,** 3461–3 (2013).

109. Uršič, T., Krivec, U., Kalan, G. & Petrovec, M. Fatal human bocavirus infection in an 18-month-old child with chronic lung disease of prematurity. *Pediatr. Infect. Dis. J.* **34,** 111–112 (2015).

110. Ziyade, N., Şirin, G., Elgörmüş, N. & Daş, T. Detection of human bocavirus DNA by multiplex PCR analysis: postmortem case report. *Balkan Med. J.* **32,** 226–9 (2015).

111. Eskola, V., Xu, M. & Söderlund-Venermo, M. Severe lower respiratory tract infection caused by human bocavirus 1 in an infant. *Pediatr. Infect. Dis. J.* **36,** 1107–1108 (2017).

112. Tabatabai, J., Fakhiri, J., Meyburg, J., Linse, K.-P., Xu, M., Söderlund-Venermo, M., Grimm, D. & Schnitzler, P. Severe human bocavirus 1 respiratory tract infection in an immunodeficient child with fatal outcome. *Pediatr. Infect. Dis. J.* **38,** e219–e222 (2019).

113. Ishiguro, T., Hirota, S., Kobayashi, Y., Takano, K., Kobayashi, Y., Shimizu, Y. & Takayanagi, N. Fatal primary human bocavirus pneumonia in an immunocompetent adult. *Intern. Med.* **59,** 421–424 (2020).

114. Ziemele, I., Xu, M., Vilmane, A., Rasa-Dzelzkaleja, S., Hedman, L., Hedman, K., Söderlund-Venermo, M., Nora-Krukle, Z., Murovska, M. & Gardovska, D. Acute human bocavirus 1 infection in child with life-threatening bilateral bronchiolitis and right-sided pneumonia: a case report. *J. Med. Case Rep.* **13,** 290 (2019).

115. Seppälä, E., Zöldi, V., Vuorinen, S., Murtopuro, S., Elonsalo, U., van Beek, J., Haveri, A., Kontio, M., Savolainen-Kopra, C., Puumalainen, T. & Sane, J. A cluster of measles linked to an imported case, Finland, 2017. *Euro Surveill.* **22,** (2017).

116. Hirsch, H. H. & Steiger, J. Polyomavirus BK. *Lancet Infect. Dis.* **3,** 611–623 (2003).

117. Tan, C. S. & Koralnik, I. J. Progressive multifocal leukoencephalopathy and other disorders caused by JC virus: clinical features and pathogenesis. *Lancet Neurol.* **9,** 425–437 (2010).

118. Musiani, M., Zerbini, M., Gentilomi, G., Plazzi, M., Gallinella, G. & Venturoli, S. Parvovirus B19 clearance from peripheral blood after acute infection. *J. Infect. Dis.* **172,** 1360–1363 (1995).

119. Cassinotti, P. & Siegl, G. Quantitative evidence for persistence of human parvovirus B19 DNA in an immunocompetent individual. *Eur. J. Clin. Microbiol. Infect. Dis.* **19,** 886–887 (2000).

120. Lefrere, J.-J., Servant-Delmas, A., Candotti, D., Mariotti, M., Thomas, I., Brossard, Y., Lefrère, F., Girot, R., Allain, J.-P. & Laperche, S. Persistent B19 infection in immunocompetent individuals: implications for transfusion safety. *Blood* **106,** 2890–2895 (2005).

121. Söderlund, M., Brown, C. S., Cohen, B. J. & Hedman, K. Accurate serodiagnosis of B19 parvovirus infections by measurement of IgG avidity. *J. Infect. Dis.* **171,** 710–713 (1995).

122. Söderlund, M., Brown, C. S., Spaan, W. J. M., Hedman, L. & Hedman, K. Epitope type-specific IgG responses to capsid proteins VP1 and VP2 of human parvovirus B19. *J. Infect. Dis.* **172,** 1431–1436 (1995).

123. Kaikkonen, L., Lankinen, H., Harjunpää, I., Hokynar, K., Söderlund-Venermo, M., Oker-Blom, C., Hedman, L. & Hedman, K. Acute-phase-specific heptapeptide epitope for diagnosis of parvovirus B19 infection. *J. Clin. Microbiol.* **37,** 3952–3956 (1999).

124. Kaikkonen, L., Söderlund-Venermo, M., Brunstein, J., Schou, O., Jensen, I. P., Rousseau, S., Caul, E. O., Cohen, B., Valle, M., Hedman, L. & Hedman, K. Diagnosis of human parvovirus B19 infections by detection of epitope-type-specific VP2 IgG. *J. Med. Virol.* **64,** 360–365 (2001).

125. Juhl, D., Özdemir, M., Dreier, J., Görg, S. & Hennig, H. Look-back study on recipients of Parvovirus B19 (B19V) DNA-positive blood components. *Vox Sang.* **109,** 305–311 (2015).

126. Kleinman, S. H., Glynn, S. A., Lee, T. H., Tobler, L. H., Schlumpf, K. S., Todd, D. S., Qiao, H., Yu, M. Y. W. & Busch, M. P. A linked donor-recipient study to evaluate parvovirus B19 transmission by blood component transfusion. *Blood* **114,** 3677–3683 (2009).

127. Paloniemi, M., Lappalainen, S., Salminen, M., Kätkä, M., Kantola, K., Hedman, L., Hedman, K., Söderlund-Venermo, M., Vesikari, T., Katka, M., Kantola, K., Hedman, L., Hedman, K., Soderlund-Venermo, M. & Vesikari, T. Human bocaviruses are commonly found in stools of hospitalized children without causal association to acute gastroenteritis. *Eur. J. Pediatr.* **173,** 1051–1057 (2014).

128. Kapusinszky, B., Minor, P. & Delwart, E. Nearly constant shedding of diverse enteric viruses by two healthy infants. *J. Clin. Microbiol.* **50,** 3427–3434 (2012).

129. Yahiro, T., Wangchuk, S., Tshering, K., Bandhari, P., Zangmo, S., Dorji, T., Tshering, K., Matsumoto, T., Nishizono, A., Söderlund-Venermo, M. & Ahmed, K. Novel human bufavirus genotype 3 in children with severe diarrhea, Bhutan. *Emerg. Infect. Dis.* **20,** 1037–9 (2014).

130. Altay, A., Yahiro, T., Bozdayi, G., Matsumoto, T., Sahin, F., Ozkan, S., Nishizono, A., Söderlund-Venermo, M. & Ahmed, K. Bufavirus genotype 3 in Turkish children with severe diarrhoea. *Clin. Microbiol. Infect.* **21,** 965.e1-965.e4 (2015).

131. Väisänen, E., Fu, Y., Hedman, K. & Söderlund-Venermo, M. Human protoparvoviruses. *Viruses* **9,** (2017).

132. Ilyas, M., Mietzsch, M., Kailasan, S., Väisänen, E., Luo, M., Chipman, P., Smith, J. K., Kurian, J., Sousa, D., McKenna, R., Söderlund-Venermo, M. & Agbandje-McKenna, M. Atomic resolution structures of human bufaviruses determined by cryo-electron microscopy. *Viruses* **10,** (2018).

133. Yang, S., Liu, D., Wang, Y., Qu, F., He, Y., Sun, Z., Shen, Q., Li, W., Fu, X., Deng, X., Zhang, W. & Delwart, E. Bufavirus protoparvovirus in feces of wild rats in China. *Virus Genes* **52,** 130–133 (2016).

134. Hillis, D. M. & Bull, J. J. An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. *Syst. Biol.* **42,** 182–192 (1993).

135. Siqueira, J. D., Dominguez-Bello, M. G., Contreras, M., Lander, O., Caballero-Arias, H., Xutao, D., Noya-Alarcon, O. & Delwart, E. Complex virome in feces from Amerindian children in isolated Amazonian villages. *Nat. Commun.* **9,** 4270 (2018).

136. Tokarz, R., Hyams, J. S., Mack, D. R., Boyle, B., Griffiths, A. M., LeLeiko, N. S., Sauer, C. G., Shah, S., Markowitz, J., Baker, S. S., Rosh, J., Baldassano, R. N., Kugathasan, S., Walters, T., Tagliafierro, T., Sameroff, S., Lee, B., Che, X., Oleynik, A., Denson, L. A. & Lipkin, W. I. Characterization of stool virome in children newly diagnosed with moderate to severe ulcerative colitis. *Inflamm. Bowel Dis.* **25,** 1656–1662 (2019).

137. Fernandes, M. A., Verstraete, S. G., Phan, T. G., Deng, X., Stekol, E., LaMere, B., Lynch, S. V., Heyman, M. B. & Delwart, E. Enteric virome and bacterial microbiota in children with ulcerative colitis and Crohn disease. *J. Pediatr. Gastroenterol. Nutr.* **68,** 30–36 (2019).

138. Kluge, M., Campos, F. S., Tavares, M., de Amorim, D. B., Valdez, F. P., Giongo, A., Roehe, P. M. & Franco, A. C. Metagenomic survey of viral diversity obtained from feces of Subantarctic and South American fur seals. *PLoS One* **11,** e0151921 (2016).

139. Mollerup, S., Fridholm, H., Vinner, L., Kjartansdóttir, K. R., Friis-Nielsen, J., Asplund, M., Herrera, J. A. R., Steiniche, T., Mourier, T., Brunak, S., Willerslev, E., Izarzugaza, J. M. G., Hansen, A. J. & Nielsen, L. P. Cutavirus in cutaneous malignant melanoma. *Emerg. Infect. Dis.* **23,** 363–365 (2017).

140. Phan, T. & Nagaro, K. Cutavirus: a newly discovered parvovirus on the rise. *Infect. Genet. Evol.* **80,** 104175 (2020).

141. Kreuter, A., Nasserani, N., Tigges, C., Oellig, F., Silling, S., Akgül, B. & Wieland, U. Cutavirus infection in primary cutaneous B- and T-cell lymphoma. *JAMA Dermatology* **154,** 965 (2018).

142. Wieland, U., Silling, S., Hufbauer, M., Mauch, C., Zigrino, P., Oellig, F., Kreuter, A. & Akgül, B. No evidence for role of cutavirus in malignant melanoma. *Emerg. Infect. Dis.* **25,** 1600–16002 (2019).

143. Bergallo, M., Daprà, V., Fava, P., Ponti, R., Calvi, C., Fierro, M. T., Quaglino, P., Galliano, I. & Montanari, P. Lack of detection of cutavirus DNA using PCR real time in Cutaneous T-cell Lymphomas (CTCL). *G. Ital. Dermatol. Venereol.* **Oct 29,** (2018).

144. Marchini, A., Bonifati, S., Scott, E. M., Angelova, A. L. & Rommelaere, J. Oncolytic parvoviruses: from basic virology to clinical applications. *Virol. J.* **12,** 6 (2015).

145. Bretscher, C. & Marchini, A. H-1 parvovirus as a cancer-killing agent: past, present, and future. *Viruses* **11,** 562 (2019).

146. Handley, S. A., Thackray, L. B., Zhao, G., Presti, R., Miller, A. D., Droit, L., Abbink, P., Maxfield, L. F., Kambal, A., Duan, E., Stanley, K., Kramer, J., Macri, S. C., Permar, S. R., Schmitz, J. E., Mansfield, K., Brenchley, J. M., Veazey, R. S., Stappenbeck, T. S., Wang, D., Barouch, D. H. & Virgin, H. W. Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. *Cell* **151,** 253–66 (2012).

147. Sasaki, M., Orba, Y., Anindita, P. D., Ishii, A., Ueno, K., Hang'ombe, B. M., Mweene, A. S., Ito, K. & Sawa, H. Distinct lineages of bufavirus in wild shrews and nonhuman primates. *Emerg. Infect. Dis.* **21,** 1230–3 (2015).

148. Kemenesi, G., Dallos, B., Görföl, T., Estók, P., Boldogh, S., Kurucz, K., Oldal, M., Marton, S., Bányai, K. & Jakab, F. Genetic diversity and recombination within bufaviruses: Detection of a novel strain in Hungarian bats. *Infect. Genet. Evol.* **33,** 288–292 (2015).

149. Wu, Z., Yang, L., Ren, X., He, G., Zhang, J., Yang, J., Qian, Z., Dong, J., Sun, L., Zhu, Y., Du, J., Yang, F., Zhang, S. & Jin, Q. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J.* **10,** 609–620 (2016).

150. Sasaki, M., Gonzalez, G., Wada, Y., Setiyono, A., Handharyani, E., Rahmadani, I., Taha, S., Adiani, S., Latief, M., Kholilullah, Z. A., Subangkit, M., Kobayashi, S., Nakamura, I., Kimura, T., Orba, Y., Ito, K. & Sawa, H. Divergent bufavirus harboured in megabats represents a new lineage of parvoviruses. *Sci. Rep.* **6,** 24257 (2016).

151. Hargitai, R., Pankovics, P., Kertész, A. M., Bíró, H., Boros, Á., Phan, T. G., Delwart, E. & Reuter, G. Detection and genetic characterization of a novel parvovirus distantly related to human bufavirus in domestic pigs. *Arch. Virol.* **161,** 1033–1037 (2016).

152. Liu, L., Schwarz, L., Ullman, K., Ahola, H., Qiu, Y., Ma, Z. & Hennig-Pauka, I. Identification of a novel bufavirus in domestic pigs by a viral metagenomic approach. *J. Gen. Virol.* **97,** 1592–1596 (2016).

153. Sun, Y.-K., Chen, Y.-J., Cai, Y., Zhu, D.-H., Pan, H.-M., Wei, Y.-F., Han, X.-L., Ji, C.-H., Lu, G., Wang, H., Ma, C.-Q. & Zhang, G.-H. First report and genetic diversity of porcine bufavirus in China. *Virol. J.* **17,** 2 (2020).

154. Siqueira, J. D., Ng, T. F., Miller, M., Li, L., Deng, X., Dodd, E., Batac, F. & Delwart, E. Endemic infection of stranded southern sea otters (Enhydra lutris nereis) with novel parvovirus, polyomavirus, and adenovirus. *J. Wildl. Dis.* **53,** 532–542 (2017).

155. Martella, V., Lanave, G., Mihalov-Kovács, E., Marton, S., Varga-Kugler, R., Kaszab, E., Di Martino, B., Camero, M., Decaro, N., Buonavoglia, C. & Bányai, K. Novel parvovirus related to primate bufaviruses in dogs. *Emerg. Infect. Dis.* **24,** 1061–1068 (2018).

156. Li, J., Cui, L., Deng, X., Yu, X., Zhang, Z., Yang, Z., Delwart, E., Zhang, W. & Hua, X. Canine bufavirus in faeces and plasma of dogs with diarrhoea, China. *Emerg. Microbes Infect.* **8,** 245–247 (2019).

157. Sun, W., Zhang, S., Huang, H., Wang, W., Cao, L., Zheng, M., Yin, Y., Zhang, H., Lu, H. & Jin, N. First identification of a novel parvovirus distantly related to human bufavirus from diarrheal dogs in China. *Virus Res.* **265,** 127–131 (2019).

158. Diakoudi, G., Lanave, G., Capozza, P., Di Profio, F., Melegari, I., Di Martino, B., Pennisi, M. G., Elia, G., Cavalli, A., Tempesta, M., Camero, M., Buonavoglia, C., Bányai, K. & Martella, V. Identification of a novel parvovirus in domestic cats. *Vet. Microbiol.* **228,** 246–251 (2019).

159. Melegari, I., Di Profio, F., Palombieri, A., Sarchese, V., Diakoudi, G., Robetto, S., Orusa, R., Marsilio, F., Bányai, K., Martella, V. & Di Martino, B. Molecular detection of canine bufaviruses in wild canids. *Arch. Virol.* **164,** 2315–2320 (2019).

160. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41,** 95–98 (1999).

161. Altan, E., Aiemjoy, K., Phan, T. G., Deng, X., Aragie, S., Tadesse, Z., Callahan, K. E., Keenan, J. & Delwart, E. Enteric virome of Ethiopian children participating in a clean water intervention trial. *PLoS One* **13,** e0202054 (2018).

162. Aiemjoy, K., Altan, E., Aragie, S., Fry, D. M., Phan, T. G., Deng, X., Chanyalew, M., Tadesse, Z., Callahan, E. K., Delwart, E. & Keenan, J. D. Viral species richness and composition in young children with loose or watery stool in Ethiopia. *BMC Infect. Dis.* **19,** 53 (2019).

163. Altay Koçak, A., Öcal, M., Polat, M., Kanık Yüksek, S., Aktaş Tapısız, A., Tezer, H., Özkul, A., Ergünay, K., Bozdayı, G. & Ahmed, K. [Multicenter investigation of bufavirus in the etiology of viral central nervous system infections of adults and children]. *Mikrobiyol. Bul.* **51,** 191–194 (2017).

164. Dickinson, A., Xu, M., Silén, S., Wang, Y., Fu, Y., Sadeghi, M., Toppinen, M., Carpén, T., Hedman, K., Mäkitie, A. & Söderlund-Venermo, M. Newly detected DNA viruses in juvenile nasopharyngeal angiofibroma (JNA) and oral and oropharyngeal squamous cell carcinoma (OSCC/OPSCC). *Eur. Arch. Otorhinolaryngol.* **276,** 613–617 (2019).

165. Okitsu, S., Khamrin, P., Takanashi, S., Thongprachum, A., Hoque, S. A., Takeuchi, H., Khan, M. A., Hasan, S. M. T., Iwata, T., Shimizu, H., Jimba, M., Hayakawa, S., Maneekarn, N. & Ushijima, H. Molecular detection of enteric viruses in the stool samples of children without diarrhea in Bangladesh. *Infect. Genet. Evol.* **77,** 104055 (2020).

166. Martin, E. T., Fairchok, M. P., Kuypers, J., Magaret, A., Zerr, D. M., Wald, A. & Englund, J. A. Frequent and prolonged shedding of bocavirus in young children attending daycare. *J. Infect. Dis.* **201,** 1625–32 (2010).

167. Smits, S. L., Schapendonk, C. M. E., van Beek, J., Vennema, H., Schürch, A. C., Schipper, D., Bodewes, R., Haagmans, B. L., Osterhaus, A. D. M. E. & Koopmans, M. P. New viruses in idiopathic human diarrhea cases, the Netherlands. *Emerg. Infect. Dis.* **20,** 1218–22 (2014).

168. Chieochansin, T., Vutithanachot, V., Theamboonlers, A. & Poovorawan, Y. Bufavirus in fecal specimens of patients with and without diarrhea in Thailand. *Arch. Virol.* **160,** 1781–1784 (2015).

169. Huang, D.-D., Wang, W., Lu, Q.-B., Zhao, J., Guo, C.-T., Wang, H.-Y., Zhang, X.-A., Tong, Y.-G., Liu, W. & Cao, W.-C. Identification of bufavirus-1 and bufavirus-3 in feces of patients with acute diarrhea, China. *Sci. Rep.* **5,** 13272 (2015).

170. Ayouni, S., Estienney, M., Hammami, S., Neji Guediche, M., Pothier, P., Aouni, M., Belliot, G. & de Rougemont, A. Cosavirus, salivirus and bufavirus in diarrheal Tunisian infants. *PLoS One* **11,** e0162255 (2016).

171. Altan, E., Del Valle Mendoza, J., Deng, X., Phan, T. G., Sadeghi, M. & Delwart, E. L. Small circular rep-encoding single-stranded DNA genomes in Peruvian diarrhea virome. *Genome Announc.* **5,** e00822-17 (2017).

172. Thongprachum, A., Fujimoto, T., Takanashi, S., Saito, H., Okitsu, S., Shimizu, H., Khamrin, P., Maneekarn, N., Hayakawa, S. & Ushijima, H. Detection of nineteen enteric viruses in raw sewage in Japan. *Infect. Genet. Evol.* **63,** 17–23 (2018).

173. Guerrero-Latorre, L., Romero, B., Bonifaz, E., Timoneda, N., Rusiñol, M., Girones, R. & Rios-Touma, B. Quito's virome: Metagenomic analysis of viral diversity in urban streams of Ecuador's capital city. *Sci. Total Environ.* **645,** 1334–1343 (2018).

174. Nantachit, N., Khamrin, P., Kumthip, K., Malasao, R. & Maneekarn, N. Molecular surveillance and genetic analyses of bufavirus in environmental water in Thailand. *Infect. Genet. Evol.* **75,** 104013 (2019).

175. Mohammad, H. A., Madi, N. M. & Al-Nakib, W. Analysis of viral diversity in stool samples from infants and children with acute gastroenteritis in Kuwait using Metagenomics approach. *Virol. J.* **17,** 10 (2020).

176. Salimans, M. M. M., Van Bussel, M. J. A. W. M., Brown, C. S. & Spaan, W. J. M. Recombinant parvovirus B19 capsids as a new substrate for detection of B19-specific IgG and IgM antibodies by an enzyme-linked immunosorbent assay. *J. Virol. Methods* **39,** 247–258 (1992).

177. Kantola, K., Sadeghi, M., Antikainen, J., Kirveskari, J., Delwart, E., Hedman, K. & Söderlund-Venermo, M. Real-time quantitative PCR detection of four human bocaviruses. *J. Clin. Microbiol.* **48,** 4044–4050 (2010).

178. Sharp, C. P., Lail, A., Donfield, S., Simmons, R., Leen, C., Klenerman, P., Delwart, E., Gomperts, E. D. & Simmonds, P. High frequencies of exposure to the novel human parvovirus PARV4 in hemophiliacs and injection drug users, as detected by a serological assay for PARV4 antibodies. *J. Infect. Dis.* **200,** 1119–25 (2009).

179. Lahtinen, A., Kivelä, P., Hedman, L., Kumar, A., Kantele, A., Lappalainen, M., Liitsola, K., Ristola, M., Delwart, E., Sharp, C., Simmonds, P., Söderlund-Venermo, M. & Hedman, K. Serodiagnosis of primary infections with human parvovirus 4, Finland. *Emerg. Infect. Dis.* **17,** 79–82 (2011).

180. Simmons, R., Sharp, C., McClure, C. P., Rohrbach, J., Kovari, H., Frangou, E., Simmonds, P., Irving, W., Rauch, A., Bowness, P., Klenerman, P. & Swiss HIV Cohort Study. Parvovirus 4 infection and clinical outcome in high-risk populations. *J. Infect. Dis.* **205,** 1816–20 (2012).

181. von Linstow, M.-L., Rosenfeldt, V., Lindberg, E., Jensen, L., Hedman, L., Li, X., Väisänen, E., Hedman, K. & Norja, P. Absence of novel human parvovirus (PARV4) in Danish mothers and children. *J. Clin. Virol.* **65,** (2015).

182. Maple, P. A., Beard, S., Parry, R. P. & Brown, K. E. Testing UK blood donors for exposure to human parvovirus 4 using a time-resolved fluorescence immunoassay to screen sera and Western blot to confirm reactive samples. *Transfusion* **53,** 2575–84 (2013).

183. Sharp, C. P., Vermeulen, M., Nébié, Y., Djoko, C. F., LeBreton, M., Tamoufe, U., Rimoin, A. W., Kayembe, P. K., Carr, J. K., Servant-Delmas, A., Laperche, S., Harrison, G. L. A., Pybus, O. G., Delwart, E., Wolfe, N. D., Saville, A., Lefrère, J. J. & Simmonds, P. Epidemiology of human parvovirus 4 infection in sub-Saharan Africa. *Emerg. Infect. Dis.* **16,** 1605–7 (2010).

184. Panning, M., Kobbe, R., Vollbach, S., Drexler, J. F., Adjei, S., Adjei, O., Drosten, C., May, J. & Eis-Hübinger, A. M. Novel human parvovirus 4 genotype 3 in infants, Ghana. *Emerg. Infect. Dis.* **16,** 1143–6 (2010).

185. Drexler, J. F., Reber, U., Muth, D., Herzog, P., Annan, A., Ebach, F., Sarpong, N., Acquah, S., Adlkofer, J., Adu-Sarkodie, Y., Panning, M., Tannich, E., May, J., Drosten, C. & Eis-Hübinger, A. M. Human parvovirus 4 in nasal and fecal specimens from children, Ghana. *Emerg. Infect. Dis.* **18,** 1650–3 (2012).

186. May, J., Drexler, J. F., Reber, U., Sarpong, N., Adjei, O., Panning, M., Drosten, C. & Eis-Hübinger, A. M. Human parvovirus 4 viremia in young children, Ghana. *Emerg. Infect. Dis.* **18,** 1690–2 (2012).

187. Matthews, P. C., Sharp, C. P., Malik, A., Gregory, W. F., Adland, E., Jooste, P., Goulder, P. J. R., Simmonds, P. & Klenerman, P. Human parvovirus 4 infection among mothers and children in South Africa. *Emerg. Infect. Dis.* **21,** 713–5 (2015).

188. Sharp, C. P., Gregory, W. F., Hattingh, L., Malik, A., Adland, E., Daniels, S., Zyl, A. van, Carlson, J. M., Wareing, S., Ogwu, A., Shapiro, R., Riddell, L., Chen, F., Ndung'u, T., Goulder, P. J. R., Klenerman, P., Simmonds, P., Jooste, P. & Matthews, P. C. PARV4 prevalence, phylogeny, immunology and coinfection with HIV, HBV and HCV in a multicentre African cohort. *Wellcome Open Res.* **2,** (2017).

189. Lavoie, M., Sharp, C. P., Pépin, J., Pennington, C., Foupouapouognigni, Y., Pybus, O. G., Njouom, R. & Simmonds, P. Human parvovirus 4 infection, Cameroon. *Emerg. Infect. Dis.* **18,** 680–3 (2012).

190. Willemze, R., Cerroni, L., Kempf, W., Berti, E., Facchetti, F., Swerdlow, S. H. & Jaffe, E. S. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. *Blood* **133,** 1703–1714 (2019).

191. Mirvish, J. J., Pomerantz, R. G., Falo, L. D. & Geskin, L. J. Role of infectious agents in cutaneous T-cell lymphoma: Facts and controversies. *Clin. Dermatol.* **31,** 423–431 (2013).

192. Mollerup, S., Asplund, M., Friis-Nielsen, J., Kjartansdóttir, K. R., Fridholm, H., Hansen, T. A., Herrera, J. A. R., Barnes, C. J., Jensen, R. H., Richter, S. R., Nielsen, I. B., Pietroni, C., Alquezar-Planas, D. E., Rey-Iglesia, A., Olsen, P. V. S., Rajpert-De Meyts, E., Groth-Pedersen, L., von Buchwald, C., Jensen, D. H., Gniadecki, R., Høgdall, E., Langhoff, J. L., Pete, I., Vereczkey, I., Baranyai, Z., Dybkaer, K., Johnsen, H. E., Steiniche, T., Hokland, P., Rosenberg, J., Baandrup, U., Sicheritz-Pontén, T., Willerslev, E., Brunak, S., Lund, O., Mourier, T., Vinner, L., Izarzugaza, J. M. G., Nielsen, L. P. & Hansen, A. J. High-throughput sequencing-based investigation of viruses in human cancers by multienrichment approach. *J. Infect. Dis.* **220,** 1312–1324 (2019).

193. Kreuter, A., Pantelaki, I., Michalowitz, A.-L., Wieland, U., Cerroni, L., Oellig, F. & Tigges, C. CD30-positive primary cutaneous anaplastic large cell lymphoma with coexistent pseudocarcinomatous hyperplasia. *Clin. Exp. Dermatol.* **43,** 585–588 (2018).

194. Wieland, U., Silling, S., Scola, N., Potthoff, A., Gambichler, T., Brockmeyer, N. H., Pfister, H. & Kreuter, A. Merkel cell polyomavirus infection in HIV-positive men. *Arch. Dermatol.* **147,** 401 (2011).